



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 165311**

**TO: Medina A Ibrahim**  
**Location: REM/2B07/2C18**  
**Art Unit: 1638**

*Sept 14*, 2005

**Case Serial Number: 10/780002**

**From: P. Sheppard**  
**Location: Remsen Building**  
**Phone: (571) 272-2529**

**sheppard@uspto.gov**

### **Search Notes**

**This Page Blank (uspto)**

From: Ibrahim, Medina A.  
Sent: Friday, September 09, 2005 12:30 PM  
To: STIC-Biotech/ChemLib  
Subject: 10/780,002

Please search the following:

1. SEQ ID NO: 1 and 36
2. DNA encoding SEQ ID NO: 2.
3. an oligo of SEQ ID NO: 1 or 36.

Please search both commercial and patent (including pending) databases. Thanks

Medina A. Ibrahim  
Patent Examiner, GAU-1638  
transgenic plants and plant breeding  
Remsen-2B07  
mailbox-2C18  
(571) 272-0797

RECEIVED  
SEP - 9 2005  
STIC-BIOTECH/CHEM. DIVISION  
(STIC)

\*\*\*\*\*

STAFF USE ONLY

Searcher: \_\_\_\_\_  
Searcher Phone: 2-\_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Rev. Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search

NA#: \_\_\_\_\_ AA#: \_\_\_\_\_  
Interference: \_\_\_\_\_ SPDI: \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure#: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_

**This Page Blank (uspto)**



Db 35 GGTTCGATCAGTGAAGAGATTACATTGTGTCAC 71  
|||||

RESULT 3  
AAQ61367/c  
ID AAQ61367 standard; DNA; 258 BP.

XX AAQ61367;  
XX  
XX 25-MAR-2003 (revised)  
XX 16-MAR-1994 (first entry)  
XX Human brain Expressed Sequence Tag EST01360.

XX Gene transcription product; genetic markers; tagging; in vivo;  
XX transcription; mapping; locations; chromosomes; chromosomal; ss.

XX Homo sapiens.  
XX W09316178-A2.  
XX  
XX 19-AUG-1993.  
XX  
XX 12-FEB-1993; 93WO-US001294.  
XX  
XX 12-FEB-1992; 92US-00837195.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICE.  
XX  
XX Venter CU, Adams MD, Moreno RF;  
XX  
XX WPI; 1993-272882/34.  
XX

XX Enriched oligonucleotides and corresp. sequences - used as markers for  
XX human genes transcribed in-vivo, facilitate tagging of most human genes.  
XX  
XX Example 4; Page 480; 500pp; English.  
XX

XX The Expressed Sequence Tag was isolated from a human brain cDNA library  
XX as part of a large set of ESTs which can be used as markers for human  
XX genes transcribed in vivo. They can be used to facilitate tagging of most  
XX human genes, for mapping locations of expressed genes on chromosomes, for  
XX individual or forensic identification, for mapping locations of disease-  
XX associated genes, for identification of tissue type, and for prep. of  
XX antisense sequences, probes and constructs. EST01360 has a "poor" coding  
XX probability as evaluated using the coding-region prediction program CRM.  
XX See also AAQ59041-061440. (Updated on 25-MAR-2003 to correct PN field.)  
XX (Updated on 25-MAR-2003 to correct PF field.)

XX Sequence 258 BP; 54 A; 58 C; 75 G; 68 T; 0 U; 3 Other;

Query Match 2.5%; Score 27; DB 2; Length 258;  
Best Local Similarity 100.0%; Pred. No. 3.1;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1049 CTTTCCTTTTGAAGAGAGATTACATTGTGTCAC 1075  
|||||  
DB 136 CTTTCCTTTTGAAGAGAGATTACATTGTGTCAC 110

RESULT 4  
ABX54351/c  
ID ABX54351 standard; cDNA; 321 BP.

XX ABX54351;  
XX  
XX 25-FEB-2003 (first entry)  
XX  
XX Bovine EST associated with lactation/muscle/fat deposition #4280.  
XX  
XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;  
XX muscle deposition; fat deposition; genome mapping; gene identification;

KW gene analysis; cattle breeding.

XX Bos Taurus.

XX US2002137160-A1.

XX 26-SEP-2002.

XX 26-OCT-2001; 2001US-00983965.

XX 17-DEC-1998; 98US-0113678P.

XX 15-DEC-1999; 99US-00465231.

XX (BYAT/) BYATT J C.

XX (MATH/) MATHIALAGAN N.

XX (TAON/) TAO N.

XX (WARR/) WARREN W C.

XX Byatt JC, Mathialagan N, Tao N, Warren WC;

XX WPI; 2003-102386/09.

XX Purified nucleic acid molecules, useful for genome mapping, gene  
XX identification and analysis, cattle breeding or preparation of constructs  
XX for cattle gene expression and genetically improved cattle.

XX Claim 2; SEQ ID NO 4280; 38pp; English.

XX The invention relates to a purified nucleic acid molecule associated with  
XX lactation or muscle and fat deposition (designated LMFD), derived from  
XX cattle, and the LMFD nucleic acid can specifically hybridise to a second  
XX nucleic acid molecule comprising any of 5912 nucleotide sequences,  
XX appearing as ABX50072-ABX55983, or complements of them. Also included are  
XX (1) a transformed cell having a nucleic acid comprising an LMFD nucleic  
XX acid linked to a promoter and a 3' non-translated sequence that  
XX functions in the cell to cause termination of transcription and addition  
XX of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and  
XX (2) determining a level or pattern of a molecule in a bovine cell or  
XX tissue comprising: (a) incubating a marker nucleic acid (comprising any  
XX of the 5912 nucleic acid sequences or its complement or fragment) with a  
XX complementary nucleic acid molecule obtained from the bovine cell or  
XX tissue, where hybridisation between the marker nucleic acid and the  
XX complementary nucleic acid permits the detection of the molecule; and (b)  
XX detecting the level or pattern of the complementary nucleic acid, where  
XX the detection of the complementary nucleic acid is predictive of the  
XX level or pattern of the molecule. The LMFD nucleic acid is used for  
XX determining a level or pattern of a molecule in a bovine cell or tissue.  
XX It is useful for genome mapping, gene identification and analysis, cattle  
XX breeding, preparation of constructs for use in cattle gene expression, or  
XX for genetically improving cattle. The present sequence is one of the 5912  
XX bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The present  
XX sequence was not shown in the specification but was obtained in  
XX electronic format from the USPTO web site:  
XX seqdata.uspto.gov/sequence.html?DocID=20020137160

XX Sequence 321 BP; 92 A; 72 C; 57 G; 100 T; 0 U; 0 Other;

Query Match 2.5%; Score 27; DB 8; Length 321;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CCTTTTGGAAAAAAGAAAAAAGAAAAA 1079  
|||||  
DB 47 CCTTTTGGAAAAAAGAAAAAAGAAAAA 21

RESULT 5

ABL62609/c

ID ABL62609 standard; DNA; 446 BP.

XX ABL62609;

XX 15-MAY-2002 (first entry)

XX DE Colon adenocarcinoma related gene sequence SEQ ID NO: 946.  
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX Homo sapiens.  
XX WO2001194629-A2.  
XX 13-DEC-2001.  
XX 30-MAY-2001; 2001WO-US010838.  
XX 05-JUN-2000; 2000US-0209473P.  
PR 05-JUN-2000; 2000US-0209531P.  
PR 18-SEP-2000; 2000US-0233133P.  
PR 18-SEP-2000; 2000US-0233617P.  
PR 20-SEP-2000; 2000US-0234009P.  
PR 20-SEP-2000; 2000US-0234034P.  
PR 20-SEP-2000; 2000US-0234052P.  
PR 22-SEP-2000; 2000US-0234509P.  
PR 22-SEP-2000; 2000US-0234567P.  
PR 25-SEP-2000; 2000US-0234923P.  
PR 25-SEP-2000; 2000US-0234924P.  
PR 25-SEP-2000; 2000US-0235077P.  
PR 25-SEP-2000; 2000US-0235082P.  
PR 25-SEP-2000; 2000US-0235134P.  
PR 25-SEP-2000; 2000US-0235280P.  
PR 26-SEP-2000; 2000US-0235637P.  
PR 26-SEP-2000; 2000US-0235638P.  
PR 27-SEP-2000; 2000US-0235711P.  
PR 27-SEP-2000; 2000US-0235720P.  
PR 27-SEP-2000; 2000US-0235840P.  
PR 27-SEP-2000; 2000US-0235863P.  
PR 28-SEP-2000; 2000US-0236028P.  
PR 28-SEP-2000; 2000US-0236032P.  
PR 28-SEP-2000; 2000US-0236033P.  
PR 28-SEP-2000; 2000US-0236034P.  
PR 28-SEP-2000; 2000US-0236109P.  
PR 28-SEP-2000; 2000US-0236111P.  
PR 29-SEP-2000; 2000US-0236842P.  
PR 29-SEP-2000; 2000US-0236891P.  
PR 02-OCT-2000; 2000US-0237172P.  
PR 02-OCT-2000; 2000US-0237173P.  
PR 02-OCT-2000; 2000US-0237278P.  
PR 02-OCT-2000; 2000US-0237294P.  
PR 02-OCT-2000; 2000US-0237295P.  
PR 02-OCT-2000; 2000US-0237316P.  
PR 03-OCT-2000; 2000US-0237425P.  
PR 03-OCT-2000; 2000US-0237598P.  
PR 03-OCT-2000; 2000US-0237604P.  
PR 03-OCT-2000; 2000US-0237606P.  
PR 03-OCT-2000; 2000US-0237608P.  
PR 01-NOV-2000; 2000US-0244867P.  
PR 01-NOV-2000; 2000US-0245084P.  
XX (AVAL-) AVALON PHARM.  
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Soppet DR, Weaver Z;  
XX WPI; 2002-188264/24.  
XX Screening for anti-neoplastic agent involves exposing cells to a chemical  
PT agent to be tested for anti-neoplastic activity, and determining a change  
PT in expression of a gene of a signature gene set.  
XX Claim 1; SEQ ID NO 946; 44pp; English.  
XX The present invention describes a method (M1) for screening for an anti-

CC neoplastic agent. The method involves exposing cells to a chemical agent  
CC to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABU61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening an  
CC anti-neoplastic agent, and can be used for producing a product which is  
CC the data collected with respect to the anti-neoplastic agent as a result  
CC of M1, and the data is sufficient to convey the chemical structure and/or  
CC properties of the agent. M1 can be used in the treatment of cancer such  
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's  
CC tumour  
XX  
SQ Sequence 446 BP; 125 A; 97 C; 84 G; 140 T; 0 U; 0 Other;  
Query Match 2.5%; Score 27; DB 6; Length 446;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1053 CCTTTTTCAGAAAAA 1079  
Db 29 CCTTTTTCAGAAAAA 3  
RESULT 6  
ABZ73132  
ID ABZ73132 standard; cDNA; 552 BP.  
XX AC ABZ73132;  
XX 10-APR-2003 (first entry)  
XX Rice leaf EST, SEQ ID NO:17.  
XX Rice; leaf; EST; expressed sequence tag; plant; biochip; DNA array;  
KW cloning; detection; heterosis; hybrid vigour; transgene detection;  
KW herbicide screening; pesticide screening; disease diagnosis;  
XX medical research; agriculture; ss.  
OS Oryza sativa.  
XX CN1364937-A.  
XX 21-AUG-2002.  
XX 31-OCT-2001; 2001CN-00137675.  
XX 31-OCT-2001; 2001CN-00137675.  
XX (UYZH-) UNIV ZHEJIANG.  
XX Li D, Dong H;  
XX WPI; 2003-000551/01.  
XX Rice leaf specific expression sequence label and its gene chip.  
XX Claim 1; Page 16 (Disclosure); 30pp; Chinese.  
XX The invention relates to 50 novel ESTs (expressed sequence tags; ABZ73116  
CC -ABZ73165) obtained from a rice leaf cDNA library. The invention also  
CC relates to a biochip comprising these ESTs. The biochip of the invention  
CC can be used in a variety of agricultural applications. It can be used in  
CC the cloning of genes which confer useful characteristics in crop species,  
CC to predict hybrid vigour (heterosis) at an early stage, in the detection  
CC of transgenic agricultural produce, in screening for novel herbicides and  
CC pesticides and in disease diagnosis. The biochip may also be used in  
CC medical research. The present sequence represents a rice leaf EST of the  
CC invention

```
XX SQ Sequence 552 BP; 135 A; 160 C; 105 G; 152 T; 0 U; 0 Other;
Query Match 2.5%; Score 27; DB 8; Length 552;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1051 TTCTTTTGAAGAAAAA 1077
DB 526 TTCTTTTGAAGAAAAA 552

RESULT 7
AAF16319
ID AAF16319 standard; cDNA; 1795 BP.
XX
AC AAF16319;
XX
XX 13-MAR-2001 (first entry)
XX Human prostate cancer antigen nucleotide sequence SEQ ID NO:754.
XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
XX neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
XX vulnery; gastrointestinal; nephrotropic; antiinfective; gynaecological;
XX antibacterial; gene therapy; neural; immune; reproductive; renal;
XX gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX wound; infectious disease; ss.
XX Homo sapiens.
XX WO200055174-A1.
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US005988.
XX
XX 12-MAR-1999; 99US-0124270P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
XX
XX Rosen CA, Ruben SM;
XX WPI; 2000-587513/55.
XX P-PSDB; AAB57116.
XX
XX Prostate cancer associated gene sequences, referred to as prostate cancer
XX antigens, useful for treatment, prevention, and diagnosis of disorders
XX such as prostate cancer.
XX Claim 1; Page 1174-1175; 2338pp; English.
XX
XX AAF15566 to AAF16505 encode the human prostate cancer associated
XX proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
XX The prostate cancer antigens can have neuroprotective, cytostatic,
XX cardioactive, immunomodulatory, muscular, vulnery, gastrointestinal,
XX nephrotropic, antiinfective, gynaecological and antibacterial activities,
XX and can be used in gene therapy. The prostate cancer antigen
XX polynucleotides may be used for detection of prostate cancer, chromosome
XX identification, as chromosome markers, and for numerous other diagnostic
XX or research purposes. The prostate cancer antigens may be used to treat
XX disorders such as neural, immune, muscular, reproductive,
XX gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
XX AAB57303 represent sequences used in the exemplification of the present
XX invention
XX
XX Sequence 1795 BP; 473 A; 375 C; 447 G; 497 T; 0 U; 3 Other;

Query Match 2.5%; Score 27; DB 3; Length 1795;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1053 CCTTTTGAAGAAAAA 1079
DB 1757 CCTTTTGAAGAAAAA 1783

RESULT 8
ABL90691
ID ABL90691 standard; cDNA; 1795 BP.
XX
AC ABL90691;
XX
XX 24-MAY-2002 (first entry)
XX Human polynucleotide SEQ ID NO 1253.
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
XX vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein; gene; ss.
XX Homo sapiens.
XX WO200190304-A2.
XX 29-NOV-2001.
XX
XX 18-MAY-2001; 2001WO-US016450.
XX
XX 19-MAY-2000; 2000US-0205515P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Birse CE, Rosen CA;
XX WPI; 2002-122018/16.
XX P-PSDB; ABB90282.
XX
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
XX prevention of neural, immune system, muscular, cardiovascular, renal and proliferative
XX disorders.
XX Claim 4; SEQ ID NO 1253; 2081pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
XX (ABB9040-ABB90444) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
XX breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
XX cardiovascular disorders such as myocardial ischaemias; (d) wound healing
XX ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
XX infectious diseases such as viral, bacterial, fungal and parasitic
XX infections. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1795 BP; 473 A; 376 C; 447 G; 497 T; 0 U; 2 Other;

Query Match 2.5%; Score 27; DB 6; Length 1795;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CCTTTTGAAGAAAAA 1079
DB 1757 CCTTTTGAAGAAAAA 1783
```

CC or systemic cytokine damage, to stimulate chondrocyte growth, to prevent  
 CC skin aging due to sunburn, to change a mammal's mental state or physical  
 CC state by influencing biorhythms, cardiac rhythms, depression, memory,  
 CC stress and for accelerating wound healing. (I), (II) and/or their agonist  
 CC or antagonist are useful as food additives or preservatives to increase  
 CC or decrease storage capabilities, fat content, lipid, protein,  
 CC carbohydrate, vitamin, mineral or other nutritional components. (I) is  
 CC useful for screening therapeutic compounds. (II) is useful in forensic  
 CC biology for detecting DNA sequences and as diagnostic probes for  
 CC detecting the presence of specific mRNA in a particular cell type  
 XX  
 XX SQ Sequence 342 BP; 116 A; 77 C; 59 G; 90 T; 0 U; 0 Other;

Query Match 2.4%; Score 26; DB 3; Length 342;  
 Best Local Similarity 100.0%; Pred. No. 6.9;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 CTTTGTGAAAAA 1079  
 DB 316 CTTTGTGAAAAA 341

RESULT 10  
 AAC79873  
 ID AAC79873 standard; cDNA; 375 BP.  
 XX  
 AC AAC79873;  
 XX  
 DT 09-FEB-2001 (first entry)  
 XX  
 DE Human secreted protein encoding cDNA for gene 25.  
 XX  
 KW Human; secreted protein; cytostatic; antiarthritic; antiasthmatic;  
 KW immunosuppressive; antiarteriosclerotic; antiinflammatory; nootropic;  
 KW neuroprotective; antidiabetic; tranquiliser; vulnery; antibacterial;  
 KW antipsoriatic; antiarrhythmic; antirheumatic; cardiant; anti-HIV;  
 KW autoimmune disorder; allergic condition; cardiovascular disorder; cancer;  
 KW neurological disease; tissue repair; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200055176-A2.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 09-MAR-2000; 2000WO-US006057.  
 XX  
 PR 12-MAR-1999; 99US-0124142P.  
 PR 11-JUN-1999; 99US-0138597P.  
 PR 03-DEC-1999; 99US-0168666P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Ruben SM, Komatsoulis G;  
 XX  
 DR WPI: 2000-638176/61.  
 DR P-PSDB; AAB44860.  
 XX  
 PT Novel 49 human secreted proteins useful for diagnosis, prevention and  
 PT treatment of disorders including neurological, cell proliferative,  
 PT cardiovascular, and autoimmune/inflammatory disorders and microbial  
 PT infections.  
 XX  
 PS Claim 1a; Page 351; 405pp; English.

This invention describes a novel isolated polypeptide (I) comprising an  
 CC amino acid sequence at least 95 % identical to a polypeptide sequence  
 CC selected from 49 polypeptides encoded by polynucleotide sequences  
 CC included in American Type Culture Collection (ATCC) deposit number  
 CC 203917, defined in the specification. The products of the invention have  
 CC cytostatic, antiarthritic, antiasthmatic, immunosuppressive, nootropic,  
 CC antiarteriosclerotic, antiinflammatory, neuroprotective, antidiabetic,  
 CC tranquiliser, vulnery, antibacterial, antipsoriatic, antiarrhythmic,  
 CC antirheumatic, cardiant and anti-HIV activity. (I) or a nucleic acid (II)  
 CC encoding (I) is useful for preventing, treating or ameliorating a medical  
 CC condition and for diagnosing a pathological condition or susceptibility  
 CC to the condition. (I) is useful for identifying a binding partner which  
 CC affects the activity of the polypeptide and for identifying an activity  
 CC in a biological sample. (I), (II) or an antibody (IV) specific to (I) is  
 CC also useful for treating or preventing a disease, disorder or condition  
 CC associated with aberrant expression of (I). Diseases treated or diagnosed  
 CC include immune disorders such as autoimmune diseases, blood protein  
 CC disorders, anemia, allergic reactions and conditions such as asthma,  
 CC organ rejection or graft-versus-host disease, inflammation, hyper  
 CC proliferative disorders, cardiovascular disorders such as arterioarterial  
 CC fistula, arrhythmias, arteriosclerosis, coronary thrombosis, organ  
 CC regeneration, cancer, neovascular glaucoma, diabetic retinopathy,  
 CC rheumatoid arthritis, psoriasis, diseases associated with increased  
 CC apoptosis that include acquired immunodeficiency syndrome (AIDS),  
 CC neurological diseases such as Parkinson's disease, viral, bacterial,  
 CC fungal or parasitic diseases. They are also used to repair, replace or  
 CC protect tissue damage by congenital defects, to treat trauma, in surgery,  
 CC including cosmetic plastic surgery, to treat fibrosis, reperfusion injury

CC antirheumatic, cardiant and anti-HIV activity. (I) or a nucleic acid (II) encoding (I) is useful for preventing, treating or ameliorating a medical condition and for diagnosing a pathological condition or susceptibility to the condition. (I) is useful for identifying a binding partner which affects the activity of the polypeptide and for identifying an activity in a biological sample. (I), (II) or an antibody (IV) specific to (I) is also useful for treating or preventing a disease, disorder or condition CC associated with aberrant expression of (I). Diseases treated or diagnosed CC include immune disorders such as autoimmune diseases, blood protein CC disorders, anemia, allergic reactions and conditions such as asthma, CC organ rejection or graft-versus-host disease, inflammation, hyper CC proliferative disorders, cardiovascular disorders such as arterioarterial CC fistula, arrhythmias, arteriosclerosis, coronary thrombosis, organ CC regeneration, cancer, neovascular glaucoma, diabetic retinopathy, CC rheumatoid arthritis, psoriasis, diseases associated with increased CC apoptosis that include acquired immunodeficiency syndrome (AIDS), CC neurological diseases such as Parkinson's disease, viral, bacterial, CC fungal or parasitic diseases. They are also used to repair, replace or CC protect tissue damage by congenital defects, to treat trauma, in surgery, CC including cosmetic plastic surgery, to treat fibrosis, reperfusion injury CC or systemic cytokine damage, to stimulate chondrocyte growth, to prevent CC skin aging due to sunburn, to change a mammal's mental state or physical CC state by influencing biorhythms, cardiac rhythms, depression, memory, CC stress and for accelerating wound healing. (I), (II) and/or their agonist CC or antagonist are useful as food additives or preservatives to increase CC or decrease storage capabilities, fat content, lipid, protein, CC carbohydrate, vitamin, mineral or other nutritional components. (I) is CC useful for screening therapeutic compounds. (II) is useful in forensic CC biology for detecting DNA sequences and as diagnostic probes for CC detecting the presence of specific mRNA in a particular cell type

XX  
SQ Sequence 375 BP; 111 A; 89 C; 74 G; 96 T; 0 U; 5 Other;

Query Match 2.4%; Score 26; DB 3; Length 375;  
Best Local Similarity 100.0%; Pred. No. 6.8;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 CTTTGTGAAAAAAAAAAAAAAAAA 1079  
DB 299 CTTTGTGAAAAAAAAAAAAAAAAA 324

## RESULT 11

AAF26589

ID AAF26589 standard; DNA; 375 BP.

AC AAF26589;

XX 27-MAR-2001 (first entry)

XX DNA encoding human secreted protein #43.

XX Secreted protein; gene therapy; vaccine; cancer; leukemia;  
KW autoimmune disease; allergy; inflammation; graft rejection;  
KW hyperproliferation; cardiovascular; infection; ss.

OS Homo sapiens.

XX WO200076531-A1.

PN 21-DEC-2000.

XX 01-JUN-2000; 2000WO-US015137.

XX 11-JUN-1999; 99US-0138625P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis GA;

XX WPI; 2001-071148/08.

XX Nucleic acids encoding 47 human secreted polypeptides, useful for

PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease and diabetic retinopathy.

XX Claim 1; Page 469; 525pp; English.

XX The present invention relates to 26 secreted human proteins. The proteins may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate polypeptide expression. For example, they may be used in gene therapy or in vaccines. Typical of diseases which are potentially treatable are cancers (including leukemia), autoimmune diseases, allergies, inflammation, graft rejection, hyperproliferation, cardiovascular diseases (particularly critical limb ischemia and coronary disease) and any involving abnormal angiogenesis, neurodegeneration and/or infectious diseases

SQ Sequence 375 BP; 111 A; 89 C; 74 G; 96 T; 0 U; 5 Other;

Query Match 2.4%; Score 26; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 6.8;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 CTTTGTGAAAAAAAAAAAAAAAAA 1079  
DB 299 CTTTGTGAAAAAAAAAAAAAAAAA 324

## RESULT 12

AAS11362

ID AAS11362 standard; DNA; 814 BP.

AC AAS11362;

XX 24-OCT-2001 (first entry)

XX Soybean DNA encoding invertase inhibitor #2.

DE Soybean; invertase inhibitor; ds; gene therapy; environmental stress;  
KW kernel development; antisense.

OS Glycine max.

XX Key Location/Qualifiers

FH CDS 99..638

FT /\*tag= a

FT /product= "Invertase inhibitor"

FT /note= "This sequence is specifically claimed in claim 2"

XX WO200158939-A2.

XX 16-AUG-2001.

XX 12-FEB-2001; 2001WO-US004492.

XX 10-FEB-2000; 2000US-0181509P.

XX (PION-) PIONEER HI-BRED INT INC.

XX (DUPO ) DU PONT DE NEMOURS &amp; CO E I.

XX Helentjaris T, Bate NJ, Allen SM;

XX WPI; 2001-502706/55.

XX P-PSDB; AAU05777.

PT An isolated polypeptide when recombinantly expressed in a plant is useful for modulating invertase activity and increasing yield in the plant.

XX Claim 2; Page 67-68; 83pp; English.

XX The invention relates to novel plant invertase inhibitors or a yeast invertase which, when recombinantly expressed in a plant, can modulate invertase activity and increase yield in the plant. Chimeric invertase inhibitors are useful for modulating invertase activity and increasing yield in a plant (especially crop species) when used to transform the

```

CC plant and are also useful for modulating kernel development and
CC protecting plants against the harmful/detrimental effects of stress and
CC adverse environmental conditions. Yeast invertase is less sensitive to
CC invertase inhibitors therefore is an attractive option to supplement
CC invertase activity in a plant using gene therapy. The novel invertase
CC inhibitor nucleic acids may be used in their antisense form. The present
CC sequence encodes a soybean invertase inhibitor
XX
SQ Sequence 814 BP; 241 A; 183 C; 138 G; 252 T; 0 U; 0 Other;
Query Match 2.4%; Score 26; DB 5; Length 814;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1054 CTTTGTGAAAAAAAAAAAAAAAAAAAA 1079
DB 770 CTTTGTGAAAAAAAAAAAAAAAAAAAA 795
RESULT 13
ADAI13386
ID ADAI13386 standard; cDNA; 1099 BP.
AC ADAI13386;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human intracellular signalling molecule INTSIG-23 cDNA, SEQ ID NO:68.
XX
KW Human; intracellular signalling molecule; INTSIG;
KW cell proliferative disorder; cancer; atherosclerosis;
KW autoimmune disorder; inflammatory disorder; infection;
KW neurological disorder; developmental disorder; endocrine disorder;
KW cytostatic; antiarteriosclerotic; nootropic; neuroprotective;
KW cerebroprotective; immunosuppressive; antiinflammatory; gene therapy;
KW gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 98..625
FT /tag= a
FT /product= "INTSIG-23"
XX
FN WO2003031568-A2.
XX
PD 17-APR-2003.
XX
PE 16-AUG-2002; 2002WO-US026322.
XX
PR 17-AUG-2001; 2001US-0313245P.
PR 24-AUG-2001; 2001US-0314751P.
PR 31-AUG-2001; 2001US-0316752P.
PR 31-AUG-2001; 2001US-0316847P.
PR 14-SEP-2001; 2001US-0322188P.
PR 28-SEP-2001; 2001US-0326390P.
PR 12-OCT-2001; 2001US-0328952P.
PR 19-OCT-2001; 2001US-0345468P.
PR 12-APR-2002; 2002US-0372499P.
XX
PA (INCYT-) INCYTE GENOMICS INC.
XX
PI Yue H, Lu DAM, Swarnakar A, Tang YT, Griffin JA, Emerling BM;
PI Forsythe IJ, Yao MG, Ramkumar J, Richardson TW, Becha SD, Lee EA;
PI Warren BA, Lehr-Nason PM, Baughn MR, Li JX, Duggan BM, Gietzen KJ;
PI Lal PG, Borowsky ML, Ison CH, Thangavelu K, Xu Y, Lee S;
PI Elliott VS, Sprague WW, Azimzai Y, Hafalia AJA, Ding L, Nguyen DB;
PI Honchell CD, Luo W, Chawla NK, Marquis JP, Jackson JL, Tran UK;
XX
WPI; 2003-393436/37.
DR P-PSDB; ADAI13341.
XX
XX New human intracellular signalling molecules (INTSIG)), useful for

```

---

```

PT diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant INTSIG expression e.g. cancer, diabetes, epilepsy, or
PT infections.
XX
PS Claim 5; Page 339; 367pp; English.
XX
CC The invention relates to 45 human intracellular signalling molecules
CC (INTSIG), INTSIG-1 to INTSIG-45, and the cDNAs encoding them (ADAI13319-
CC ADAI13408). The invention also encompasses expression constructs, host
CC cells and transgenic organisms comprising an INTSIG nucleic acid sequence
CC ; the recombinant preparation of a INTSIG; an antibody against a INTSIG;
CC methods of detection of INTSIG proteins or nucleic acids; a micro-array
CC containing INTSIG nucleic acids; methods of screening compounds for their
CC ability to modulate INTSIG activity or expression; and pharmaceutical
CC compositions comprising an INTSIG protein, an INTSIG antibody, an INTSIG
CC agonist or INTSIG antagonist. The INTSIG proteins, nucleic acids or
CC compositions comprising them are useful in diagnosing, treating or
CC preventing a variety of disorders, including cell proliferative disorders
CC (e.g., cancer or atherosclerosis); autoimmune/inflammatory disorders
CC (e.g., AIDS, allergies, anaemia, asthma, bronchitis, gout, multiple
CC sclerosis, osteoarthritis, rheumatoid arthritis, or systemic lupus
CC erythematosus); neurological disorders (e.g., epilepsy, stroke,
CC Alzheimer's disease, dementia, or Parkinson's disease); viral, bacterial,
CC fungal, parasitic, protozoan or helminthic infections; neurological
CC disorders (e.g., epilepsy, stroke, Alzheimer's disease, dementia, or
CC Parkinson's disease); developmental disorders (e.g., achondroplastic
CC dwarfism, Cushing's disease, hypothyroidism or hydrocephalus); endocrine
CC disorders (e.g., disorders of the hypothalamus or pituitary gland or
CC diabetes); reproductive disorders; or vesicle-trafficking disorders
CC (e.g., cystic fibrosis or hypercholesterolaemia). The present sequence
CC represents an INTSIG cDNA of the invention.
XX
SQ Sequence 1099 BP; 397 A; 166 C; 227 G; 309 T; 0 U; 0 Other;
Query Match 2.4%; Score 26; DB 10; Length 1099;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1054 CTTTGTGAAAAAAAAAAAAAAAAAAAA 1079
DB 1069 CTTTGTGAAAAAAAAAAAAAAAAAAAA 1094
RESULT 14
AAQ73389
ID AAQ73389 standard; cDNA to mRNA; 1664 BP.
XX
AC AAQ73389;
XX
DT 25-MAR-2003 (revised)
DT 09-JUN-1995 (first entry)
XX
DE Rabies virus glycoprotein coding sequence.
XX
KW Rabies virus; glycoprotein; cancer; vaccine; colonic cancer; melanoma;
KW breast cancer; kidney; tumour cell; fibroblast; transfect;
KW transformed cells; cytokine; co-stimulating molecule; culture;
KW heat stable antigen; ss.
XX
OS Rabies virus.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..6
FT /tag= a
FT CDS 7..1581
FT /tag= b
FT sig_peptide 7..63
FT /tag= c
FT mat_peptide 64..1578
FT /tag= d
XX
XX WO9421808-A1.
XX

```

PD 29-SEP-1994.  
XX  
PF 18-MAR-1994; 94WO-EP000859.  
XX  
PR 19-MAR-1993; 93AT-00000556.  
PR 10-AUG-1993; 93DE-04326821.  
XX  
XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
PA (GETH ) GENENTECH INC.  
XX  
XX Birnstiel ML, Buschle M, Cotten M, Maas G, Plank C, Schaffner G;  
PI Schmidt W, Wagner E, Zatloukal K;  
PI Schmidt W, Wagner E, Zatloukal K;  
XX  
XX WPI; 1994-317028/39.  
DR P-PSDB; AAR62438.  
XX  
XX New cancer vaccine based on autologous, transfected tumour cells -  
PT expressing immunostimulating polypeptide, for treating esp. melanoma and  
PT colonic carcinoma, also new transfection complex and transfected cells.  
XX  
XX Disclosure; Page 108-110; 160pp; German.  
XX  
XX The sequence encoding the rabies virus glycoprotein may be used in the  
CC construction of a recombinant DNA molecule which can then be used to  
CC transfect cultured tumour cells or fibroblasts. The resulting cells may  
CC be used as a cancer vaccine (the recombinant DNA molecule may  
CC alternatively comprise at least one cytokine encoding sequence or a  
CC sequence encoding a co-stimulating molecule e.g. a heat stable antigen).  
CC The transfected cells are inactivated so that they retain the ability to  
CC express the immunostimulating polypeptide (in this case the rabies  
CC glycoprotein) but they cannot replicate. In the case of transfected  
CC fibroblasts, these are combined with inactivated but not transfected  
CC tumour cells. The transfected cells provide 30 times greater expression  
CC of the immunogenic polypeptide than known vectors and the localised  
CC production avoids problems associated with systemic toxicity. Fibroblasts  
CC need not be autologous (eliminating the need to produce individual  
CC cultures) and their use allows vaccines to be made with only a small  
CC number of tumour cells. The vaccines are especially used to treat  
CC melanoma and colonic carcinoma but may also be used against e.g. kidney  
CC and breast cancers. The recombinant molecules provide one component of a  
CC three component electrically neutral complex. The other components are:  
CC (1) a conjugate between a DNA binding compound and an endosomolytic agent  
CC ; and (2) a DNA binding compound pref. conjugated to an internalisation  
CC factor that binds to a surface molecule on the cell being transfected.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
XX Sequence 1664 BP; 457 A; 381 C; 408 G; 418 T; 0 U; 0 Other;  
SQ

Query Match 2.4%; Score 26; DB 2; Length 1664;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 CTTTTTGAIAAAAAAAAAAAAAAAAAAAAA 1079  
|||||  
DB 1639 CTTTTTGAIAAAAAAAAAAAAAAAAAAAAA 1664

RESULT 15  
AAZ06371  
ID AAZ06371 standard; DNA; 1870 BP.  
XX  
XX AAZ06371;  
AC  
XX  
DT 26-OCT-1999 (first entry)  
XX  
DE Human receptor protein (HURP) 6 nucleotide sequence.  
XX  
KW receptor; cancer; autoimmune disorder; inflammation; antagonist;  
KW cell surface protein; cell signalling; antibody; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers

CDS 116..931  
FT /\*tag= a  
FT /product= "Human receptor protein 6"  
XX  
XX WO9941375-A2.  
XX 19-AUG-1999.  
XX 05-FEB-1999; 99WO-US002572.  
XX 12-FEB-1998; 98US-00022939.  
XX (INCY-) INCYTE PHARM INC.  
XX Hillman JL, Bandman O, Yue H, Guegler KJ, Corley NC, Au-Young J;  
PI Tang YT, Shah P, Lal P, Baughn M;  
XX  
XX WPI; 1999-494536/41.  
DR P-PSDB; AAY15227.  
XX  
XX New human receptor proteins, used e.g. to treat, prevent and diagnose  
PT gastrointestinal and developmental disorders - and related nucleic acids,  
PT vectors, transformed cells, antibodies, agonists and antagonists.  
XX  
XX Claim 11; Page 90; 94pp; English.  
XX This is the consensus sequence of the Human receptor protein 6 (HURP-6)  
CC derived from the following overlapping and/or extended nucleic acid  
CC sequences; incyte clones 1261646 (SYNORAT05), 1419057 (KIDNNOT09),  
CC 1651262 (PROSTUT08), 1704186 (DUODNOT02), and 3394566 (LUNGNOT28). HURP-6  
CC is expressed in cancerous, inflamed, reproductive, and gastrointestinal  
CC tissue. Therefore, HURP-6 appears to play a role in cancer,  
CC autoimmune/inflammatory disorders, reproductive disorders, and  
CC gastrointestinal disorders. This gives rise to the possibility of using  
CC an antagonist or an antibody of HURP to treat or prevent cancer or  
CC autoimmune/inflammatory disorders  
XX  
XX Sequence 1870 BP; 475 A; 402 C; 476 G; 517 T; 0 U; 0 Other;  
SQ

Query Match 2.4%; Score 26; DB 2; Length 1870;  
Best Local Similarity 100.0%; Pred. No. 5.1;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CTTTTTGAIAAAAAAAAAAAAAAAAAAAAA 1078  
|||||  
DB 1837 CTTTTTGAIAAAAAAAAAAAAAAAAAAAAA 1862

Search completed: September 12, 2005, 13:13:11  
Job time : 598.713 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: September 12, 2005, 06:38:26 ; Search time 4681.81 Seconds  
(without alignments)  
11167.303 Million call updates/sec

Title: US-10-780-002-1  
Perfect score: 1079  
Sequence: 1 acgcgggaagaagaaaaa.....gaaaaaagaaaaa 1079

Scoring table: OLIGO NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sta:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID       | Description        |
|------------|-------|-------------|--------|----|----------|--------------------|
| 1          | 1079  | 100.0       | 1079   | 8  | AY485932 | AY485932 Nicotiana |
| 2          | 37    | 3.4         | 121    | 6  | CQ089241 | CQ089241 Sequence  |
| 3          | 30    | 2.8         | 2810   | 10 | BC080732 | BC080732 Mus muscu |
| 4          | 30    | 2.8         | 12336  | 3  | AC116550 | AC116550 Dictyoste |
| 5          | 30    | 2.8         | 199593 | 10 | AL671871 | AL671871 Mouse DNA |
| 6          | 30    | 2.8         | 250078 | 3  | AE014829 | AE014829 Plasmodiu |
| 7          | 30    | 2.8         | 252632 | 3  | AE014818 | AE014818 Plasmodiu |
| 8          | 29    | 2.7         | 250029 | 3  | AE014820 | AE014820 Plasmodiu |
| 9          | 28    | 2.6         | 118189 | 2  | AL513496 | AL513496 Homo sapi |
| 10         | 28    | 2.6         | 145638 | 9  | HS357K22 | AL022720 Human DNA |
| 11         | 28    | 2.6         | 224697 | 2  | AC135583 | AC135583 Rattus no |
| 12         | 28    | 2.6         | 230038 | 2  | AC109654 | AC109654 Rattus no |
| 13         | 28    | 2.6         | 257686 | 2  | AC121225 | AC121225 Rattus no |
| 14         | 27    | 2.5         | 207    | 6  | CQ672612 | CQ672612 Sequence  |
| 15         | 27    | 2.5         | 446    | 6  | AX330437 | AX330437 Sequence  |
| 16         | 27    | 2.5         | 1778   | 9  | BC065299 | BC065299 Homo sapi |
| 17         | 27    | 2.5         | 1830   | 9  | BC063001 | BC063001 Homo sapi |
| 18         | 27    | 2.5         | 138461 | 2  | AC113576 | AC113576 Tetraodon |
| 19         | 27    | 2.5         | 156881 | 5  | BX088545 | BX088545 Zebrafish |

|    |    |     |        |    |           |                      |
|----|----|-----|--------|----|-----------|----------------------|
| 20 | 27 | 2.5 | 162712 | 2  | AC015677  | AC015677 Homo sapi   |
| 21 | 27 | 2.5 | 179391 | 2  | AP001283  | AP001283 Homo sapi   |
| 22 | 27 | 2.5 | 206736 | 2  | AP001200  | AP001200 Homo sapi   |
| 23 | 26 | 2.4 | 332    | 6  | CQ709902  | CQ709902 Sequence    |
| 24 | 26 | 2.4 | 814    | 6  | AR491517  | AR491517 Sequence    |
| 25 | 26 | 2.4 | 814    | 6  | AX214366  | AX214366 Sequence    |
| 26 | 26 | 2.4 | 1197   | 9  | AB046068  | AB046068 Macaca fa   |
| 27 | 26 | 2.4 | 1234   | 9  | AF038169  | AF038169 Homo sapi   |
| 28 | 26 | 2.4 | 1495   | 9  | BC063688  | BC063688 Homo sapi   |
| 29 | 26 | 2.4 | 1622   | 5  | BC061308  | BC061308 Xenopus t   |
| 30 | 26 | 2.4 | 1664   | 6  | A40148    | A40148 Sequence 1    |
| 31 | 26 | 2.4 | 1839   | 10 | AY138960  | AY138960 Mus muscu   |
| 32 | 26 | 2.4 | 1860   | 10 | BC027433  | BC027433 Mus muscu   |
| 33 | 26 | 2.4 | 1870   | 6  | BD131320  | BD131320 Human rec   |
| 34 | 26 | 2.4 | 2013   | 8  | AY735703  | AY735703 Arabidops   |
| 35 | 26 | 2.4 | 2080   | 8  | PTV13772  | PTV13772 Populus tri |
| 36 | 26 | 2.4 | 2359   | 9  | BC050626  | BC050626 Homo sapi   |
| 37 | 26 | 2.4 | 5966   | 9  | HS0808857 | BS4648706 Homo sapi  |
| 38 | 26 | 2.4 | 24479  | 3  | AC114265  | AC114265 Dictyoste   |
| 39 | 26 | 2.4 | 24739  | 9  | AY621074  | AY621074 Homo sapi   |
| 40 | 26 | 2.4 | 41721  | 9  | AC145291  | AC145291 Homo sapi   |
| 41 | 26 | 2.4 | 41823  | 9  | AC092428  | AC092428 Homo sapi   |
| 42 | 26 | 2.4 | 67322  | 2  | AC102007  | AC102007 Mus muscu   |
| 43 | 26 | 2.4 | 91692  | 9  | AC084253  | AC084253 Homo sapi   |
| 44 | 26 | 2.4 | 98345  | 6  | CQ870165  | CQ870165 Sequence    |
| 45 | 26 | 2.4 | 113153 | 5  | BX005344  | BX005344 Zebrafish   |

ALIGNMENTS

|            |  |             |      |        |                 |
|------------|--|-------------|------|--------|-----------------|
| RESULT 1   | AY485932   | 1079 bp     | mRNA | linear | PLN 07-JAN-2004 |
| LOCUS      | Nicotiana tabacum salicylic acid-binding protein 2 mRNA, complete cds.   |             |      |        |                 |
| DEFINITION | AY485932   |             |      |        |                 |
| ACCESSION  | AY485932   |             |      |        |                 |
| VERSION    | AY485932.1   | GI:40549302 |      |        |                 |
| KEYWORDS   | Nicotiana tabacum (common tobacco)   |             |      |        |                 |
| SOURCE     | Nicotiana tabacum  |             |      |        |                 |
| ORGANISM   | Nicotiana tabacum  |             |      |        |                 |
| REFERENCE  | 1 (bases 1 to 1079)  |             |      |        |                 |
| AUTHORS    | Kumar,D. and Klessig,D.F.  |             |      |        |                 |
| TITLE      | High-affinity salicylic acid-binding protein 2 is required for plant innate immunity and has salicylic acid-stimulated lipase activity |             |      |        |                 |
| JOURNAL    | Proc. Natl. Acad. Sci. U.S.A. 100 (26), 16101-16106 (2003)   |             |      |        |                 |
| PUBMED     | 14673096   |             |      |        |                 |
| REFERENCE  | 2 (bases 1 to 1079)  |             |      |        |                 |
| AUTHORS    | Kumar,D. and Klessig,D.F.  |             |      |        |                 |
| TITLE      | Direct Submission  |             |      |        |                 |
| JOURNAL    | Submitted (24-NOV-2003) BTI, Tower Rd., Ithaca, NY 14853, USA  |             |      |        |                 |
| FEATURES   | Location/Qualifiers  |             |      |        |                 |
| Source     | 1..1079  |             |      |        |                 |
|            | /organism="Nicotiana tabacum"  |             |      |        |                 |
|            | /mol_type="mRNA"   |             |      |        |                 |
|            | /cultivar="Xanthi nc [NN]"   |             |      |        |                 |
|            | /db_xref="taxon:4097"  |             |      |        |                 |
|            | 42..824  |             |      |        |                 |
|            | /note="lipase; SAMP2"  |             |      |        |                 |
| CDS        | /codon_start=1   |             |      |        |                 |
|            | /product="salicylic acid-binding protein 2"  |             |      |        |                 |
|            | /protein_id="AA087711.1"   |             |      |        |                 |
|            | /db_xref="GI:40549303"   |             |      |        |                 |
|            | /translation="MKEGKHFLVHGACHGMSWYKLPLEAGHKVTLDAASGT  |             |      |        |                 |
|            | DLRKIEELRTLYDVTLPMEWMESLSADEKVLVGHSLGGMNLGAMWEKYPKIVAAV  |             |      |        |                 |
|            | FLAAMPDSVHNSFVLQYNERTPAENWLTQFLPYGSPPEPLSMFGPKFLAHKL   |             |      |        |                 |
|            | YQLCPDIALALASLVRPFSSFMEDLSKAKYFTDRFGSVKRVYLIIVCTEDRGIPREFQ   |             |      |        |                 |
|            | RWQIDNIGVTAIEIKGADHMAMLCEPKQKLCASLLEIAHKYN"  |             |      |        |                 |
| ORIGIN     |  |             |      |        |                 |

```
Query Match      100.0%; Score 1079; DB 8; Length 1079;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1079; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1  ACGCGGGGAAAGAAAGAACTAAACAAGCATATAAATTCAAATGAAGGAAGGAAACACT 60
DB 1  ACGCGGGGAAAGAAAGAACTAAACAAGCATATAAATTCAAATGAAGGAAGGAAACACT 60

QY 61  TTGTTTTAGTACATGTCATGCCATGCCATGGAGTTGGAGTTGGTACAAGCTTAAGCCACTGC 120
DB 61  TTGTTTTAGTACATGTCATGCCATGCCATGGAGTTGGAGTTGGTACAAGCTTAAGCCACTGC 120

QY 121  TAGAAGCTCAGGCCATAGAGTTACAGCCCTTGATTTAGCAGCTTCTGGCAGCTGATTTGA 180
DB 121  TAGAAGCTCAGGCCATAGAGTTACAGCCCTTGATTTAGCAGCTTCTGGCAGCTGATTTGA 180

QY 181  GAAATATAGAGGCTTCGCACACTTTATGATTATATCTTTGGCATTGATGGAGTTGATGG 240
DB 181  GAAATATAGAGGCTTCGCACACTTTATGATTATATCTTTGGCATTGATGGAGTTGATGG 240

QY 241  AATCTCTTTTCAGCAGATGAGAGGTTATATTAGTGGGCATAGTCTTGTGGTATGAAT 300
DB 241  AATCTCTTTTCAGCAGATGAGAGGTTATATTAGTGGGCATAGTCTTGTGGTATGAAT 300

QY 301  TGGGACTTGTATGGAAGATATCCAAAGATCTATGCTGCTGTTTCTTGGCTGCTT 360
DB 301  TGGGACTTGTATGGAAGATATCCAAAGATCTATGCTGCTGTTTCTTGGCTGCTT 360

QY 361  TCATGCTCATGTTCTGTTACAACTCCTCTTTGTTTGGAAACAGATATATGACGAGCC 420
DB 361  TCATGCTCATGTTCTGTTACAACTCCTCTTTGTTTGGAAACAGATATATGACGAGCC 420

QY 421  CAGCCGAGAAATTTGGTTGGATCTCAGTTTTTACCATAATGTTTCCCTGAAGGCCACTGA 480
DB 421  CAGCCGAGAAATTTGGTTGGATCTCAGTTTTTACCATAATGTTTCCCTGAAGGCCACTGA 480

QY 481  CATCCATGTTTTTGGCCCAAGTTCTTGGCTCACAAGCTTACCAAGCTATGCTCTCTG 540
DB 481  CATCCATGTTTTTGGCCCAAGTTCTTGGCTCACAAGCTTACCAAGCTATGCTCTCTG 540

QY 541  AGGATCTTCATTTAGCATCATATGTTGTCAGACCAAGCTCTTTGTTTATGGAAGCACTAT 600
DB 541  AGGATCTTCATTTAGCATCATATGTTGTCAGACCAAGCTCTTTGTTTATGGAAGCACTAT 600

QY 601  CGAAGGCCAAGTATTTTACAGATGAACGGTTTGGATCAGTGAAGAGAGTTTACATTGTT 660
DB 601  CGAAGGCCAAGTATTTTACAGATGAACGGTTTGGATCAGTGAAGAGAGTTTACATTGTT 660

QY 661  GCACTGAGGATTAAGGCATACCGAAGAAATTCAGGATGGCCAAATTTGACAACATTGGTG 720
DB 661  GCACTGAGGATTAAGGCATACCGAAGAAATTCAGGATGGCCAAATTTGACAACATTGGTG 720

QY 721  TCACCTAAGCAATAGAGATTAAGGTTGCTGATCACAATGCAATGCTATCGGAGCCCAAA 780
DB 721  TCACCTAAGCAATAGAGATTAAGGTTGCTGATCACAATGCAATGCTATCGGAGCCCAAA 780

QY 781  AACTTTGCGCCTCTCTCTTGGAAATTTGCCATAAATACAACTGATCTCTACATTATGCT 840
DB 781  AACTTTGCGCCTCTCTCTTGGAAATTTGCCATAAATACAACTGATCTCTACATTATGCT 840

QY 841  TGTCTCATGTCAGATTTTTCAGTGCATGCTGTGTAATTTTTTTCTATTTTTTCAGCCGCC 900
DB 841  TGTCTCATGTCAGATTTTTCAGTGCATGCTGTGTAATTTTTTTCTATTTTTTCAGCCGCC 900

QY 901  ATAACTGCTCTTTCCTTATTTTAAAGGATGTCAGTAAATTTTCACTCTCTTAGTGTGGAAGGCT 960
DB 901  ATAACTGCTCTTTCCTTATTTTAAAGGATGTCAGTAAATTTTCACTCTCTTAGTGTGGAAGGCT 960

QY 961  TCCACATAAGGATGTTCTCTTTCTCATTCAAGTGTGTTTATGTTGATGATACTTAAAC 1020
DB 961  TCCACATAAGGATGTTCTCTTTCTCATTCAAGTGTGTTTATGTTGATGATACTTAAAC 1020

QY 1021  CGTATCAATCTCTGTATGAAGAACTCTTCTTCTTCTTTTGAAGGAAAAAAAAAAAA 1079
DB 1021  CGTATCAATCTCTGTATGAAGAACTCTTCTTCTTCTTCTTTTGAAGGAAAAAAAAAAAA 1079

RESULT 2
CQ809241
LOCUS      CQ809241
DEFINITION Sequence 544 from Patent WO2003097790.
ACCESSION CQ809241
VERSION    CQ809241.1
KEYWORDS   GI:47114635
SOURCE     Nicotiana tabacum (common tobacco)
ORGANISM   Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Nicotiana.
1
Inz,D.G., Goossens,A., Okeman-Caldentey,K.M., Haekkinen,S.T. and
Laakso,I.J.
Genes and uses thereof to modulate secondary metabolite
biosynthesis
Patent: WO 2003097790-A 544 27-NOV-2003;
Vlaams Interuniversitair Instituut voor Biotechnologie vzw w. (BE);
VTT Biotechnology (FI)
FEATURES
source
1. .121
/organism="Nicotiana tabacum"
/mol_type="unassigned DNA"
/db_xref="taxon:4097"
ORIGIN
Query Match      3.4%; Score 37; DB 6; Length 121;
Best Local Similarity 100.0%; Pred. No. 4.3e-08;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 628  GCTTTGGATCAGTGAAGAGAGTTTACATTGTGTGCAC 664
DB 35  GCTTTGGATCAGTGAAGAGAGTTTACATTGTGTGCAC 71

RESULT 3
CQ80732/c
LOCUS      CQ80732/c
DEFINITION Mus musculus thymocyte selection-associated HMGB box gene, mRNA
ACCESSION BC080732
VERSION    BC080732.1
KEYWORDS   MGC.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2810)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,P.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullah,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahy,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalek,U., Smallus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
```

















NEDFLLRKIKNAELIKKAKLNKNEHINIKRKTGPKPAAMIKVPPSKIITIEPLPFYEE  
Y"

repeat\_region 8467. .8550  
/rpt\_type=tandem  
repeat\_region 8473. .8539  
/rpt\_type=tandem  
repeat\_region 8477. .8537  
/rpt\_type=tandem  
repeat\_region 8723. .8844  
/rpt\_type=tandem  
repeat\_region 8723. .8840  
/rpt\_type=tandem  
repeat\_region 8782. .8844  
/rpt\_type=tandem  
repeat\_region 9019. .9059  
/rpt\_type=tandem  
repeat\_region 9019. .9058  
/rpt\_type=tandem  
repeat\_region 9133. .9189  
/rpt\_type=tandem  
repeat\_region 9221. .9254  
/rpt\_type=tandem  
repeat\_region 9604. .9637  
/rpt\_type=tandem  
repeat\_region 9668. .9723  
/rpt\_type=tandem  
repeat\_region 10036. .10073  
/rpt\_type=tandem  
repeat\_region 10445. .10472  
/rpt\_type=tandem  
repeat\_region 10511. .10551  
/rpt\_type=tandem

gene complement(<10565. .>11080)  
/locus\_tag="PF14\_0241"  
mRNA complement(<10565. .>11080)  
/locus\_tag="PF14\_0241"  
CDS complement(10565. .11080)  
/locus\_tag="PF14\_0241"  
/codon\_start=1  
/product="basic transcription factor 3b, putative"  
/protein\_id="AAN36854.1"  
/db\_xref="GI:23497309"

/translation="MEKISPEILARAKLKERMGGNLRQIGKGSGARRKIKVHKNSI  
SNEKINIILKIGASYFGVDIEICVYRTGDTFLEFKPKLCASLOSNTYIVTGKPE  
HKIDINKLEGLKGNKLDNMLEKIKNDENIKNIINKESGDTPKREDEQEAANDVPD  
LVNFEVSKE"

repeat\_region 11429. .11458  
/rpt\_type=tandem  
repeat\_region 11514. .11584  
/rpt\_type=tandem  
repeat\_region 11514. .11579

Query Match 2.7%; Score 29; DB 3; Length 250029;  
Best Local Similarity 100.0%; Pred. No. 0.0008;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1051 TTCCTTTTGAAGAAAAAAGAAAAA 1079  
|||||  
Db 42212 TTCCTTTTGAAGAAAAAAGAAAAA 42184

RESULT 9  
AL513496/C AL513496 118189 bp DNA linear HTG 10-JUL-2001  
LOCUS Homo sapiens chromosome X clone RP3-359K12.  
DEFINITION Direct Submission  
ACCESSION AL513496  
VERSION AL513496.2 GI:13121498  
KEYWORDS HTG; HTGS PHASE2; HTGS\_CANCELLED.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1

AUTHORS  
TITLE  
JOURNAL

COMMENT

Lawlor,S.  
Direct Submission  
Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
requests: clonerequest@sanger.ac.uk  
On Feb 23, 2001 this sequence version replaced gi:12733864.  
----- Genome Center  
Center: Sanger Centre  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
----- Project Information  
Center project name: dj359K12  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Sequencing vector: plasmid; L08752; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Consensus quality: 118159 bases at least Q40  
Consensus quality: 118179 bases at least Q30  
Consensus quality: 118187 bases at least Q20  
Insert size: 118189; sum-of-contigs  
Insert size: 125554; 4.1% error; agarose-fp  
Quality coverage: 9.82x in Q20 bases; sum-of-contigs Quality  
coverage: 9.24x in Q20 bases; agarose-fp  
-----

\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

FEATURES  
Source

Location/Qualifiers  
1. .118189  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="X"  
/clone="RP3-359K12"  
/clone\_lib="RPCI-3"  
misc\_feature 1. .118189  
/note="assembly\_fragment:03123"

ORIGIN

Query Match 2.6%; Score 28; DB 2; Length 118189;  
Best Local Similarity 100.0%; Pred. No. 0.0028;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1052 TCCTTTTGAAGAAAAAAGAAAAA 1079  
|||||  
Db 54570 TCCTTTTGAAGAAAAAAGAAAAA 54543

RESULT 10

HS357K22/c

LOCUS

DEFINITION

Human DNA sequence from clone RP3-357K22 on chromosome Xq27.1-27.3,  
complete sequence.

ACCESSION

AL022720

VERSION

AL022720.1 GI:3980360

KEYWORDS

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 145638)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission  
Submitted (22-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk  
On Dec 8, 1998 this sequence version replaced gi:3646072.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission



```

repeat_region 106552..106937
/note="LIMAB repeat: matches 5894. .6284 of consensus"
repeat_region 107659..107726
/note="34 copies 2 mer aa 67% conserved"
repeat_region 108859..108916
/note="29 copies 2 mer ga 74% conserved"
repeat_region 114000..114043
/note="22 copies 2 mer ct 81% conserved"
repeat_region 115870..115899
/note="15 copies 2 mer tg 90% conserved"
repeat_region 116775..117108
/note="MTT2A repeat: matches 2. .304 of consensus"
repeat_region 117092..117279
/note="LIMTA repeat: matches 263. .453 of consensus"
repeat_region 121735..121786
/note="26 copies 2 mer at 80% conserved"
repeat_region 126415..126534
/note="60 copies 2 mer ta 76% conserved"
repeat_region 131714..131786
/note="78K repeat: matches 121. .193 of consensus"
repeat_region 131823..131949
/note="78K repeat: matches 1. .125 of consensus"
repeat_region 135220..135282
/note="LIP3 repeat: matches 6084. .6146 of consensus"
repeat_region 144427..144948
/note="LIM3 repeat: matches 5605. .6133 of consensus"
repeat_region 145088..145638
/note="LIPAS repeat: matches 5596. .6145 of consensus"

```

## ORIGIN

```

Query Match      2.6%; Score 28; DB 9; Length 145638;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1052 TCCTTTTGAAGAAAAA 1079
      |||
Db 112349 TCCTTTTGAAGAAAAA 112322

```

## RESULT 11

```

AC135583          224697 bp  DNA  linear  HTG 20-NOV-2002
LOCUS             Rattus norvegicus clone CH230-219F9, *** SEQUENCING IN PROGRESS
DEFINITION        *** 2 unrounded pieces.
AC135583          AC135583
VERSION           HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS          Rattus norvegicus (Norway rat)
SOURCE            Rattus norvegicus
ORGANISM          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
                  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
                  Rattus.

```

## REFERENCE

```

AUTHORS           Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,
                  Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
                  Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
                  Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
                  Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
                  Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
                  Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
                  Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
                  Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
                  Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
                  Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
                  Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
                  Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
                  Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
                  Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
                  Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
                  Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
                  Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
                  Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
                  Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,

```

Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhenwa,L., Louleeged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mahoney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokemeloh,O., Okwuonu,G., Olarnpungoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L., Puozzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,P., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savary,G., Scherer,S., Scott,G., Shatman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wlarczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

```

Direct Submission
Unpublished
2 (bases 1 to 224697)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-OCT-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 224697)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

```

## COMMENT

On Nov 20, 2002 this sequence version replaced gi:24137403.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KBVQ
Center clone name: CH230-219F9
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 209475 bases at least Q40
Consensus quality: 212691 bases at least Q30
Consensus quality: 214935 bases at least Q20
Estimated insert size: 218919; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length

```





QY 1051 TTCCTTTTGAAGAAAAA 1078  
Db 14669 TTCCTTTTGAAGAAAAA 14696

RESULT 14  
CQ672612 207 bp DNA linear PAT 03-FEB-2004  
LOCUS  
DEFINITION Sequence 17538 from Patent WO2070737.  
ACCESSION CQ672612  
VERSION CQ672612.1 GI:42169050  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Liew,C.C., Marshall,W.E. and Zhang,H.  
TITLE Compositions and methods relating to osteoarthritis  
JOURNAL Patent: WO 02070737-A 17538 12-SEP-2002;  
Chondrogene Inc. (CA)  
FEATURES  
source 1..207  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 2.5%; Score 27; DB 6; Length 207;  
Best Local Similarity 100.0%; Pred. No. 0.011; 0; Indels 0; Gaps 0;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CCTTTTGAAGAAAAA 1079  
Db 154 CCTTTTGAAGAAAAA 180

RESULT 15  
AX330437/c 446 bp DNA linear PAT 09-JAN-2002  
LOCUS  
DEFINITION Sequence 946 from Patent WO0194629.  
ACCESSION AX330437  
VERSION AX330437.1 GI:18103415  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,  
Horrihan,S., Soppet,D.R. and Weaver,Z.  
TITLE Cancer gene determination and therapeutic screening using signature  
gene sets  
JOURNAL Patent: WO 0194629-A 946 13-DEC-2001;  
Avalon Pharmaceuticals (US)  
FEATURES  
source 1..446  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 2.5%; Score 27; DB 6; Length 446;  
Best Local Similarity 100.0%; Pred. No. 0.011; 0; Indels 0; Gaps 0;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CCTTTTGAAGAAAAA 1079  
Db 29 CCTTTTGAAGAAAAA 3

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 15:37:37 ; Search time 619.757 Seconds  
(without alignments)  
9663.988 Million cell updates/sec

Title: US-10-780-002-36

Perfect score: 912

Sequence: 1 atcagcgtattctcaacaa.....ttttccgggcaactttcatc 912

Scoring table: OLIGO\_NUC

Gapop\_60.0 , Gapext 60.0

Searched: 7351250 seqs, 3283620254 residues

Word size : 0

Total number of hits satisfying chosen parameters: 14702500

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications\_NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:\*
- 22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq:\*
- 24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description         |
|------------|-------|-------------|--------|----|---------------------|
| 1          | 912   | 100.0       | 912    | 21 | US-10-780-002-36    |
| 2          | 432   | 47.4        | 483    | 10 | US-09-770-961-854   |
| 3          | 22    | 2.4         | 2466   | 19 | US-10-437-963-45846 |
| 4          | 21    | 2.3         | 1239   | 19 | US-10-437-963-31574 |
| 5          | 20    | 2.2         | 201    | 20 | US-10-719-993-14754 |
| 6          | 20    | 2.2         | 201    | 20 | US-10-719-993-14862 |
| 7          | 20    | 2.2         | 8592   | 15 | US-10-311-455-1955  |

RESULT 1

US-10-780-002-36  
; Sequence 36, Application US/10780002  
; Publication No. US20050034196A1  
; GENERAL INFORMATION:  
; APPLICANT: Klesig, Daniel F.  
; APPLICANT: Kumar, Dhirenda  
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN  
; TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, SAMP2, AND METHODS OF USE THEREOF  
; FILE REFERENCE: 3670-FO2652WO  
; CURRENT APPLICATION NUMBER: US/10/780,002  
; PRIOR FILING DATE: 2004-02-17  
; PRIOR APPLICATION NUMBER: PCT/US02/26312  
; PRIOR FILING DATE: 2002-08-16  
; PRIOR APPLICATION NUMBER: 60/312,863  
; PRIOR FILING DATE: 2001-08-16  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 36  
; LENGTH: 912  
; TYPE: DNA  
; ORGANISM: Arabidopsis Thaliana  
US-10-780-002-36

ALIGNMENTS

|   |    |    |     |        |    |                     |                   |
|---|----|----|-----|--------|----|---------------------|-------------------|
| c | 8  | 20 | 2.2 | 366710 | 20 | US-10-719-993-6792  | Sequence 6792, Ap |
| c | 9  | 19 | 2.1 | 319    | 19 | US-10-767-701-24762 | Sequence 24762, A |
| c | 10 | 19 | 2.1 | 345    | 19 | US-10-437-963-98094 | Sequence 98094, A |
| c | 11 | 19 | 2.1 | 387    | 16 | US-10-116-712-82    | Sequence 82, Appl |
| c | 12 | 19 | 2.1 | 418    | 20 | US-10-425-115-77654 | Sequence 77654, A |
| c | 13 | 19 | 2.1 | 432    | 19 | US-10-767-795-5332  | Sequence 5332, Ap |
| c | 14 | 19 | 2.1 | 440    | 16 | US-10-116-712-643   | Sequence 643, App |
| c | 15 | 19 | 2.1 | 462    | 19 | US-10-437-963-58265 | Sequence 58265, A |
| c | 16 | 19 | 2.1 | 624    | 9  | US-09-902-941-1914  | Sequence 1914, Ap |
| c | 17 | 19 | 2.1 | 624    | 9  | US-09-849-626-1914  | Sequence 1914, Ap |
| c | 18 | 19 | 2.1 | 624    | 14 | US-10-017-754-1914  | Sequence 1914, Ap |
| c | 19 | 19 | 2.1 | 624    | 16 | US-10-113-872-1914  | Sequence 1914, Ap |
| c | 20 | 19 | 2.1 | 624    | 17 | US-10-283-017-1914  | Sequence 1914, Ap |
| c | 21 | 19 | 2.1 | 636    | 9  | US-09-736-457-1668  | Sequence 1668, Ap |
| c | 22 | 19 | 2.1 | 636    | 9  | US-09-902-941-1668  | Sequence 1668, Ap |
| c | 23 | 19 | 2.1 | 636    | 9  | US-09-849-626-1668  | Sequence 1668, Ap |
| c | 24 | 19 | 2.1 | 636    | 14 | US-10-017-754-1668  | Sequence 1668, Ap |
| c | 25 | 19 | 2.1 | 636    | 16 | US-10-113-872-1668  | Sequence 1668, Ap |
| c | 26 | 19 | 2.1 | 636    | 17 | US-10-283-017-1668  | Sequence 1668, Ap |
| c | 27 | 19 | 2.1 | 1008   | 15 | US-10-156-761-7148  | Sequence 7148, Ap |
| c | 28 | 19 | 2.1 | 1310   | 17 | US-10-264-049-218   | Sequence 218, App |
| c | 29 | 19 | 2.1 | 1633   | 9  | US-09-736-457-316   | Sequence 316, App |
| c | 30 | 19 | 2.1 | 1633   | 9  | US-09-736-457-788   | Sequence 788, App |
| c | 31 | 19 | 2.1 | 1633   | 9  | US-09-902-941-316   | Sequence 316, App |
| c | 32 | 19 | 2.1 | 1633   | 9  | US-09-902-941-788   | Sequence 788, App |
| c | 33 | 19 | 2.1 | 1633   | 9  | US-09-849-626-316   | Sequence 316, App |
| c | 34 | 19 | 2.1 | 1633   | 9  | US-09-849-626-788   | Sequence 788, App |
| c | 35 | 19 | 2.1 | 1633   | 10 | US-09-476-300-316   | Sequence 316, App |
| c | 36 | 19 | 2.1 | 1633   | 14 | US-10-017-754-316   | Sequence 316, App |
| c | 37 | 19 | 2.1 | 1633   | 14 | US-10-017-754-788   | Sequence 788, App |
| c | 38 | 19 | 2.1 | 1633   | 16 | US-10-113-872-316   | Sequence 316, App |
| c | 39 | 19 | 2.1 | 1633   | 16 | US-10-113-872-788   | Sequence 788, App |
| c | 40 | 19 | 2.1 | 1633   | 17 | US-10-283-017-316   | Sequence 316, App |
| c | 41 | 19 | 2.1 | 1633   | 17 | US-10-283-017-788   | Sequence 788, App |
| c | 42 | 19 | 2.1 | 1633   | 17 | US-10-172-118-1143  | Sequence 1143, Ap |
| c | 43 | 19 | 2.1 | 1633   | 18 | US-10-342-887-1143  | Sequence 1143, Ap |
| c | 44 | 19 | 2.1 | 1633   | 21 | US-10-848-755A-139  | Sequence 139, App |
| c | 45 | 19 | 2.1 | 1679   | 20 | US-10-817-483-17    | Sequence 17, Appl |

|                       |              |              |            |             |
|-----------------------|--------------|--------------|------------|-------------|
| Query Match           | 100.0%       | Score 912;   | DB 21;     | Length 912; |
| Best Local Similarity | 100.0%;      | Pred. No. 0; | Mismatches | 0;          |
| Matches 912;          | Conservative | 0;           | Indels     | 0;          |
| Gaps                  | 0;           |              |            |             |

QY 1 ATCAGCTATTCTCAACAGCAAGACCTTAACCAAGAAAGAACTCTACGAGACA 60  
|||||







Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 45846  
LENGTH: 2466  
TYPE: DNA  
ORGANISM: Oryza sativa  
NAME/KEY: unsure  
LOCATION: (1)..(2466)  
OTHER INFORMATION: unsure at all n locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_48771C.1  
US-10-437-963-45846

Query Match 2.4%; Score 22; DB 19; Length 2466;  
Best Local Similarity 100.0%; Pred. No. 0.95;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 176 CTCGCCGCTCCGGATCGACC 197  
DB 382 CTCGCCGCTCCGGATCGACC 403

RESULT 4  
US-10-437-963-31574  
Sequence 31574, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 31574  
LENGTH: 1239  
TYPE: DNA  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_35868C.1  
US-10-437-963-31574

Query Match 2.3%; Score 21; DB 19; Length 1239;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 611 CTATCAAGAAAGAAAGTTT 631  
DB 622 CTATCAAGAAAGAAAGTTT 642

RESULT 5  
US-10-719-993-14754  
Sequence 14754, Application US/10719993  
Publication No. US20040265849A1  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
FILE REFERENCE: CL001496  
CURRENT APPLICATION NUMBER: US/10/719,993  
CURRENT FILING DATE: 2003-11-24  
NUMBER OF SEQ ID NOS: 55342  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14754  
LENGTH: 201  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-719-993-14754

Query Match 2.2%; Score 20; DB 20; Length 201;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 499 ATTGAAGATGGGACCAAAAT 518  
DB 127 ATTGAAGATGGGACCAAAAT 146

RESULT 6  
US-10-719-993-14862  
Sequence 14862, Application US/10719993  
Publication No. US20040265849A1  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
FILE REFERENCE: CL001496  
CURRENT APPLICATION NUMBER: US/10/719,993  
CURRENT FILING DATE: 2003-11-24  
NUMBER OF SEQ ID NOS: 55342  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14862  
LENGTH: 201  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-719-993-14862

Query Match 2.2%; Score 20; DB 20; Length 201;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 499 ATTGAAGATGGGACCAAAAT 518  
DB 13 ATTGAAGATGGGACCAAAAT 32

RESULT 7  
US-10-311-455-1955/c  
Sequence 1955, Application US/10311455  
Publication No. US20030143606A1  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PIRPENBROCK, Christian  
APPLICANT: BERLIN, Kurt  
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining  
FILE REFERENCE: 5013.1014  
CURRENT APPLICATION NUMBER: US/10/311,455  
CURRENT FILING DATE: 2002-12-16  
PRIOR APPLICATION NUMBER: PCT/EP01/07537  
PRIOR FILING DATE: 2001-07-02  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30

```
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1955
; LENGTH: 8592
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1955
```

```
Query Match          2.2%; Score 20; DB 15; Length 8592;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 781 CTCCAAAACCCCAAAACTCT 800
Db 2474 CTCCAAAACCCCAAAACTCT 2455
|||||
```

```
RESULT 8
US-10-719-993-6792
; Sequence 6792, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6792
; LENGTH: 366710
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(366710)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-719-993-6792
```

```
Query Match          2.2%; Score 20; DB 20; Length 366710;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 499 ATTGAAGATGGGACCAAAAT 518
Db 184736 ATTGAAGATGGGACCAAAAT 184755
|||||
```

```
RESULT 9
US-10-767-701-24762/c
; Sequence 24762, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 24762
; LENGTH: 319
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 30941244
US-10-767-701-24762
```

```
Query Match          2.1%; Score 19; DB 19; Length 319;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 169 TGTGAACTGCGCGCTCC 187
Db 145 TGTGAACTGCGCGCTCC 127
|||||
```

```
RESULT 10
US-10-437-963-98094
; Sequence 98094, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 98094
; LENGTH: 345
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_96032C.1
US-10-437-963-98094
```

```
Query Match          2.1%; Score 19; DB 19; Length 345;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 771 TGGTGATGCTCTCCAAACC 789
Db 234 TGGTGATGCTCTCCAAACC 252
|||||
```

```
RESULT 11
US-10-116-712-82/c
; Sequence 82, Application US/10116712
; Publication No. US20030194764A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Switzer, Ann
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.568
; CURRENT APPLICATION NUMBER: US/10/116,712
; CURRENT FILING DATE: 2002-04-07
; NUMBER OF SEQ ID NOS: 670
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 387
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 320_336, 345
; OTHER INFORMATION: n = A,T,C or G
US-10-116-712-82
```

```
Query Match          2.1%; Score 19; DB 16; Length 387;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 890 CCTTTTCCGGGCAACTTT 908  
|||||  
Db 137 CCTTTTCCGGGCAACTTT 119

RESULT 12  
US-10-425-115-77654/c  
; Sequence 77654, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 77654  
; LENGTH: 418  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_170847C.1  
US-10-425-115-77654

Query Match 2.1%; Score 19; DB 20; Length 418;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 880 AAAACAAACCCCTTTTCC 898  
|||||  
Db 133 AAAACAAACCCCTTTTCC 115

RESULT 13  
US-10-767-795-5332  
; Sequence 5332, Application US/10767795  
; Publication No. US20040181830A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Zhou, Yihua  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53534)B  
; CURRENT APPLICATION NUMBER: US/10/767,795  
; CURRENT FILING DATE: 2004-01-30  
; NUMBER OF SEQ ID NOS: 117596  
; SEQ ID NO 5332  
; LENGTH: 432  
; TYPE: DNA  
; ORGANISM: Gossypium hirsutum  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3145-010-Q1-K1-E3  
US-10-767-795-5332

Query Match 2.1%; Score 19; DB 19; Length 432;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 880 AAAACAAACCCCTTTTCC 898  
|||||  
Db 1 AAAACAAACCCCTTTTCC 19

RESULT 14  
US-10-116-712-643/c  
; Sequence 643, Application US/10116712  
; Publication No. US20030194764A1

; GENERAL INFORMATION:  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Switzer, Ann  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.568  
; CURRENT APPLICATION NUMBER: US/10/116,712  
; CURRENT FILING DATE: 2002-04-07  
; NUMBER OF SEQ ID NOS: 670  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 643  
; LENGTH: 440  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-116-712-643

Query Match 2.1%; Score 19; DB 16; Length 440;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGGCAACTTT 908  
|||||  
Db 137 CCTTTTCCGGGCAACTTT 119

RESULT 15  
US-10-437-963-58265/c  
; Sequence 58265, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 58265  
; LENGTH: 462  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_5999C.1  
US-10-437-963-58265

Query Match 2.1%; Score 19; DB 19; Length 462;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 AATCCAGGCGGTGAACCC 223  
|||||  
Db 217 AATCCAGGCGGTGAACCC 199

Search completed: September 12, 2005, 20:50:01  
Job time : 623.757 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 11:35:08 ; Search time 173.147 Seconds  
(without alignments)  
8618.590 Million cell updates/sec

Title: US-10-780-002-36  
Perfect score: 912  
Sequence: 1 atcacgctattctcaacaa.....ttttcgggcaactttcatc 912

Scoring table: OLIGO NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:\*\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*\*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq:\*\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID                   | Description         |
|------------|-------|-------------|--------|----------------------|---------------------|
| 1          | 22    | 2.4         | 601    | US-09-949-016-183626 | Sequence 183626,    |
| 2          | 22    | 2.4         | 45983  | US-09-949-016-17010  | Sequence 17010, A   |
| 3          | 19    | 2.1         | 393    | US-09-540-236-856    | Sequence 856, App   |
| 4          | 19    | 2.1         | 636    | US-09-702-705-1668   | Sequence 1668, App  |
| 5          | 19    | 2.1         | 636    | US-09-736-457-1668   | Sequence 1668, App  |
| 6          | 19    | 2.1         | 636    | US-09-614-124B-1668  | Sequence 1668, App  |
| 7          | 19    | 2.1         | 636    | US-09-671-325-1668   | Sequence 1668, App  |
| 8          | 19    | 2.1         | 636    | US-09-658-824-1668   | Sequence 1668, App  |
| 9          | 19    | 2.1         | 1633   | US-09-702-705-316    | Sequence 316, App   |
| 10         | 19    | 2.1         | 1633   | US-09-702-705-788    | Sequence 788, App   |
| 11         | 19    | 2.1         | 1633   | US-09-736-457-316    | Sequence 316, App   |
| 12         | 19    | 2.1         | 1633   | US-09-736-457-788    | Sequence 788, App   |
| 13         | 19    | 2.1         | 1633   | US-09-614-124B-316   | Sequence 316, App   |
| 14         | 19    | 2.1         | 1633   | US-09-614-124B-788   | Sequence 788, App   |
| 15         | 19    | 2.1         | 1633   | US-09-671-325-316    | Sequence 316, App   |
| 16         | 19    | 2.1         | 1633   | US-09-671-325-788    | Sequence 788, App   |
| 17         | 19    | 2.1         | 1633   | US-09-589-184-316    | Sequence 316, App   |
| 18         | 19    | 2.1         | 1633   | US-09-589-184-788    | Sequence 788, App   |
| 19         | 19    | 2.1         | 1633   | US-09-658-824-316    | Sequence 316, App   |
| 20         | 19    | 2.1         | 1633   | US-09-658-824-788    | Sequence 788, App   |
| 21         | 19    | 2.1         | 1635   | US-09-949-016-4625   | Sequence 4625, App  |
| 22         | 19    | 2.1         | 9544   | US-09-949-016-16367  | Sequence 16367, A   |
| 23         | 18    | 2.0         | 333    | US-09-248-796A-2805  | Sequence 2805, App  |
| 24         | 18    | 2.0         | 366    | US-09-248-796A-11450 | Sequence 11450, App |
| 25         | 18    | 2.0         | 601    | US-09-949-016-20860  | Sequence 20860, A   |
| 26         | 18    | 2.0         | 601    | US-09-949-016-110207 | Sequence 110207, A  |
| 27         | 18    | 2.0         | 601    | US-09-949-016-198581 | Sequence 198581, A  |

ALIGNMENTS

RESULT 1  
US-09-949-016-183626  
; Sequence 183626, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 183626  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-183626

|                       |                 |                 |           |             |
|-----------------------|-----------------|-----------------|-----------|-------------|
| Query Match           | 2.4%            | Score 22;       | DB 4;     | Length 601; |
| Best Local Similarity | 100.0%          | Pred. No. 0.21; |           |             |
| Matches 22;           | Conservative 0; | Mismatches 0;   | Indels 0; | Gaps 0;     |

QY 68 AGGAAACATCACTTCGTTGTTAG 89  
|||||  
DB 512 AGGAAACATCACTTCGTTGTTAG 533

RESULT 2  
US-09-949-016-17010/c  
; Sequence 17010, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498

```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17010
; LENGTH: 45983
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..- (45983)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17010

Query Match      2.4%; Score 22; DB 4; Length 45983;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      68 AGGAACATCACTTCGTGTAG 89
Db      34394 AGGAACATCACTTCGTGTAG 34373

RESULT 3
US-09-540-236-856/c
; Sequence 856, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 856
; LENGTH: 393
; TYPE: DNA
; ORGANISM: M.catarrhalis
US-09-540-236-856

Query Match      2.1%; Score 19; DB 4; Length 393;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      785 AAACCCCAAAACTCTTTG 803
Db      272 AAACCCCAAAACTCTTTG 254

RESULT 4
US-09-702-705-1668/c
; Sequence 1668, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1668
; LENGTH: 636
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-09-702-705-1668

Query Match      2.1%; Score 19; DB 4; Length 636;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      890 CCTTTTCCGGGCAACTTT 908
Db      562 CCTTTTCCGGGCAACTTT 544

RESULT 5
US-09-736-457-1668/c
; Sequence 1668, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1668
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-457-1668

Query Match      2.1%; Score 19; DB 4; Length 636;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      890 CCTTTTCCGGGCAACTTT 908
Db      562 CCTTTTCCGGGCAACTTT 544

RESULT 6
US-09-614-124B-1668/c
; Sequence 1668, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1668
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
```

## US-09-614-124B-1668

Query Match 2.1%; Score 19; DB 4; Length 636;  
Best Local Similarity 100.0%; Pred. No. 8.6;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGGCAACTTT 908  
|||||  
Db 562 CCTTTTCCGGGCAACTTT 544

## RESULT 7

US-09-671-325-1668/c  
; Sequence 1668, Application US/09671325  
; Patent No. 6667154  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C12  
; CURRENT APPLICATION NUMBER: US/09/671,325  
; CURRENT FILING DATE: 2000-09-26  
; NUMBER OF SEQ ID NOS: 1825  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1668  
; LENGTH: 636  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-671-325-1668

Query Match 2.1%; Score 19; DB 4; Length 636;  
Best Local Similarity 100.0%; Pred. No. 8.6;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGGCAACTTT 908  
|||||  
Db 562 CCTTTTCCGGGCAACTTT 544

## RESULT 8

US-09-658-824-1668/c  
; Sequence 1668, Application US/09658824  
; Patent No. 6746846  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
; FILE REFERENCE: 210121.478C11  
; CURRENT APPLICATION NUMBER: US/09/658,824  
; CURRENT FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 1788  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1668  
; LENGTH: 636  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-658-824-1668

Query Match 2.1%; Score 19; DB 4; Length 636;  
Best Local Similarity 100.0%; Pred. No. 8.6;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGGCAACTTT 908  
|||||  
Db 562 CCTTTTCCGGGCAACTTT 544

## RESULT 9

US-09-702-705-316/c  
; Sequence 316, Application US/09702705  
; Patent No. 6504010  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C14  
; CURRENT APPLICATION NUMBER: US/09/702,705  
; CURRENT FILING DATE: 2000-10-30  
; NUMBER OF SEQ ID NOS: 1833  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 316  
; LENGTH: 1633  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-702-705-316

Query Match 2.1%; Score 19; DB 4; Length 1633;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGGCAACTTT 908  
|||||  
Db 624 CCTTTTCCGGGCAACTTT 606

## RESULT 10

US-09-702-705-788/c  
; Sequence 788, Application US/09702705  
; Patent No. 6504010  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C14  
; CURRENT APPLICATION NUMBER: US/09/702,705  
; CURRENT FILING DATE: 2000-10-30  
; NUMBER OF SEQ ID NOS: 1833  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 788  
; LENGTH: 1633  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-702-705-788

Query Match 2.1%; Score 19; DB 4; Length 1633;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGGCAACTTT 908  
|||||  
Db 624 CCTTTTCCGGGCAACTTT 606

## RESULT 11

US-09-736-457-316/c  
; Sequence 316, Application US/09736457  
; Patent No. 6509448  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Lijun  
; APPLICANT: Wang, Aijun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736,457  
; CURRENT FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 316  
; LENGTH: 1633  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-736-457-316

Query Match 2.1%; Score 19; DB 4; Length 1633;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGGCAACTTT 908  
|||||  
Db 624 CCTTTTCCGGGCAACTTT 606

## RESULT 12

US-09-736-457-788/c  
; Sequence 788, Application US/09736457  
; Patent No. 6509448  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Lijun  
; APPLICANT: Wang, Aijun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736,457  
; CURRENT FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 788  
; LENGTH: 1633  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-736-457-788

Query Match 2.1%; Score 19; DB 4; Length 1633;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGGCAACTTT 908  
|||||  
Db 624 CCTTTTCCGGGCAACTTT 606

## RESULT 13

US-09-614-124B-316/c  
; Sequence 316, Application US/09614124B  
; Patent No. 6630574  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
; FILE REFERENCE: 210121.478C9  
; CURRENT APPLICATION NUMBER: US/09/614,124B  
; CURRENT FILING DATE: 2001-07-11  
; NUMBER OF SEQ ID NOS: 1668  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 316  
; LENGTH: 1633  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-614-124B-316

Query Match 2.1%; Score 19; DB 4; Length 1633;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGGCAACTTT 908  
|||||  
Db 624 CCTTTTCCGGGCAACTTT 606

## RESULT 14

US-09-614-124B-788/c  
; Sequence 788, Application US/09614124B  
; Patent No. 6630574  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
; FILE REFERENCE: 210121.478C9  
; CURRENT APPLICATION NUMBER: US/09/614,124B  
; CURRENT FILING DATE: 2001-07-11  
; NUMBER OF SEQ ID NOS: 1668  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 788  
; LENGTH: 1633  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-614-124B-788

Query Match 2.1%; Score 19; DB 4; Length 1633;  
Best Local Similarity 100.0%; Pred. No. 8.9;



Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 890 CCTTTTTCGGGCAACTTT 908  
Db 624 CCTTTTTCGGGCAACTTT 606

RESULT 15  
US-09-671-325-316/C  
; Sequence 316, Application US/09671325  
; Patent No. 6667154  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C12  
; CURRENT APPLICATION NUMBER: US/09/671,325  
; CURRENT FILING DATE: 2000-09-26  
; NUMBER OF SEQ ID NOS: 1825  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 316  
; LENGTH: 1633  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-671-325-316

Query Match 2.1%; Score 19; DB 4; Length 1633;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 890 CCTTTTTCGGGCAACTTT 908  
Db 624 CCTTTTTCGGGCAACTTT 606

Search completed: September 12, 2005, 17:29:51  
Job time : 176.147 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 08:41:37 ; Search time 2908.23 Seconds  
(without alignments)  
11936.677 Million cell updates/sec

Title: US-10-780-002-36

Perfect score: 912

Sequence: 1 atcaactattctcaacaa.....ttttccgggcaactttcatc 912

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:\*

1: gb\_est1:.\*  
2: gb\_est2:.\*  
3: gb\_hc:.\*  
4: gb\_est3:.\*  
5: gb\_est4:.\*  
6: gb\_est5:.\*  
7: gb\_est6:.\*  
8: gb\_gsa1:.\*  
9: gb\_gsa2:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 718   | 78.7        | 865    | 3     | CNS09Z6U    |
| 2          | 628   | 68.9        | 667    | 1     | AV822206    |
| 3          | 624   | 68.4        | 625    | 1     | AV785311    |
| 4          | 554   | 60.7        | 641    | 6     | CB258477    |
| 5          | 435   | 47.7        | 537    | 1     | AV783452    |
| 6          | 421   | 46.2        | 421    | 5     | BP613009    |
| 7          | 397   | 43.5        | 572    | 1     | AV822640    |
| 8          | 374   | 41.0        | 441    | 1     | AV824147    |
| 9          | 366   | 40.1        | 366    | 1     | AV554161    |
| 10         | 341   | 37.4        | 390    | 1     | AV820476    |
| 11         | 289   | 31.7        | 480    | 1     | AA04853     |
| 12         | 279   | 30.6        | 335    | 1     | AV545429    |
| 13         | 263   | 28.8        | 382    | 5     | BP646852    |
| 14         | 259   | 28.4        | 361    | 1     | AI993246    |
| 15         | 247   | 27.1        | 308    | 1     | AV559909    |
| 16         | 232   | 25.4        | 405    | 5     | BP642613    |
| 17         | 232   | 25.4        | 425    | 5     | BP642196    |
| 18         | 217   | 23.8        | 217    | 1     | AV782897    |
| 19         | 203   | 22.3        | 351    | 5     | BP604502    |
| 20         | 167   | 18.3        | 426    | 1     | AV790585    |
| 21         | 142   | 15.6        | 296    | 9     | CNS00PYH    |
| 22         | 136   | 14.9        | 444    | 9     | BX292336    |
| 23         | 99    | 10.9        | 488    | 9     | BX292222    |
| 24         | 87    | 9.5         | 159    | 1     | AI996681    |

|   |    |    |     |      |   |          |                    |
|---|----|----|-----|------|---|----------|--------------------|
| C | 25 | 85 | 9.3 | 581  | 9 | BX292335 | BX292335 Arabidops |
|   | 26 | 50 | 5.5 | 177  | 7 | R84081   | R84081 16040 Lambd |
|   | 27 | 33 | 3.6 | 556  | 8 | BZ507980 | BZ507980 BONA086TR |
|   | 28 | 32 | 3.5 | 138  | 9 | BX891283 | BX891283 Arabidops |
| C | 29 | 25 | 2.7 | 411  | 5 | BP652964 | BP652964 BP652964  |
|   | 30 | 23 | 2.5 | 730  | 8 | BZ237627 | BZ237627 CH230-333 |
| C | 31 | 22 | 2.4 | 433  | 2 | BE328824 | BE328824 hv97e01.X |
| C | 32 | 22 | 2.4 | 698  | 9 | CL567704 | CL567704 OB_Ba003  |
|   | 33 | 22 | 2.4 | 725  | 8 | BZ050121 | BZ050121 jnr44c05  |
|   | 34 | 22 | 2.4 | 1083 | 9 | CL200471 | CL200471 ZMMBB023  |
|   | 35 | 21 | 2.3 | 418  | 6 | CB765657 | CB765657 AMGNNUC:N |
|   | 36 | 21 | 2.3 | 539  | 1 | AL921928 | AL921928 BAC21928  |
| C | 37 | 21 | 2.3 | 557  | 8 | BH822165 | BH822165 BACPF17-M |
|   | 38 | 21 | 2.3 | 605  | 2 | BF099324 | BF099324 601751709 |
| C | 39 | 21 | 2.3 | 681  | 7 | CK414051 | CK414051 AUF_IpGii |
| C | 40 | 21 | 2.3 | 691  | 7 | CK413120 | CK413120 AUF_IpGii |
|   | 41 | 21 | 2.3 | 760  | 2 | BF241640 | BF241640 601878974 |
|   | 42 | 21 | 2.3 | 822  | 5 | BQ432919 | BQ432919 AGENCOURT |
| C | 43 | 21 | 2.3 | 855  | 9 | CL851022 | CL851022 OR_CBa007 |
| C | 44 | 21 | 2.3 | 889  | 6 | CA474406 | CA474406 AGENCOURT |
|   | 45 | 21 | 2.3 | 1176 | 9 | AG340901 | AG340901 Mus muscu |

#### ALIGNMENTS

RESULT 1  
CNS09Z6U  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
JOURNAL  
COMMENT  
FEATURES  
source

CNS09Z6U 865 bp mRNA linear HTC 06-FBB-2004  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
GSLTSL782C11 of Silique of strain col-0 of Arabidopsis thaliana  
(thale cress).  
BX833804  
BX833804.1 GI:42455725  
HTC; GSLT\_cDNA.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.  
1 (bases 1 to 865)  
Castelli,V., Aury,J.M., Jallou,O., Wincker,P., Clepet,C.,  
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,  
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.  
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
Annotation  
Unpublished  
2 (bases 1 to 865)  
Genoscope.  
Direct Submission  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out  
full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C.,  
URV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis  
genome released by MIPS (Munich Information center for Protein  
Sequences). 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
length  
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.  
Location/Qualifiers  
1. 865  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strains="Col-0"  
/db\_xref="taxon:3702"

```

/clone="GSLTSL178ZC11"
/tissue_type="silique"
/plasmid="pCMVSPORT_6"
1..865
/gene="At5g10300"

ORIGIN
Query Match      78.7%; Score 718; DB 3; Length 865;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 838; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 75 ATCACTTCGTGTAGTTCACACGCTTATCATGAGCGCTTGATCTGGTACAGCTCAAGC 134
DB 1 ATCACTTCGTGTAGTTCACACGCTTATCATGAGCGCTTGATCTGGTACAGCTCAAGC 60

QY 135 CCTCTCTTGAATCAGCGCGCCACCGCTTACTGTCTGCGAACTCGCCGCCCTCCGGGATCG 194
DB 61 CCTCTCTTGAATCAGCGCGCCACCGCTTACTGTCTGCGAACTCGCCGCCCTCCGGGATCG 120

QY 195 ACCCAGCAATTCACGCGCTTGAAACCGTTCGACGAATATCTCCAAACCGTTGATCGAAA 254
DB 121 ACCCAGCAATTCACGCGCTTGAAACCGTTCGACGAATATCTCCAAACCGTTGATCGAAA 180

QY 255 CCTCAATCTCTTCCAGAGACGAGGTAATTCGTGTGGATTTCAGCTTCGGAGGCA 314
DB 181 CCTCAATCTCTTCCAGAGACGAGGTAATTCGTGTGGATTTCAGCTTCGGAGGCA 240

QY 315 TCAATATCGCTCTCGCGCGGACATATTTCCGGCGAAGATTAAGTTCTTGTTCTCTCA 374
DB 241 TCAATATCGCTCTCGCGCGGACATATTTCCGGCGAAGATTAAGTTCTTGTTCTCTCA 300

QY 375 ACGCTCTTTCGCCGACAAA-CCACGTCCTTCTCAGCTTCTCGGACAAAGTATGAG 433
DB 301 ACGCTCTTTCGCCGACAAA-CCACGTCCTTCTCAGCTTCTCGGACAAAGTATGAG 360

QY 434 ATGCTGGAGGTTGGAGATTGTGAGTTTTCATCTCATGAACAAAGAAATGGGACGATG 493
DB 361 ATGCTGGAGGTTGGAGATTGTGAGTTTTCATCTCATGAACAAAGAAATGGGACGATG 420

QY 494 AGTTTATGAAGATGGACCAAAATTCATGAAGGACGCTCTTTACCAAAATGTCCCAT 553
DB 421 AGTTTATGAAGATGGACCAAAATTCATGAAGGACGCTCTTTACCAAAATGTCCCAT 480

QY 554 GAGGATTACGAGTGGCAAAAATGTTGCATAGGCAAGGTCATTTTTCACAGAGATCTA 613
DB 481 GAGGATTACGAGTGGCAAAAATGTTGCATAGGCAAGGTCATTTTTCACAGAGATCTA 540

QY 614 TCAAGAAAGAAAGTTTAGCGGAAGGATATGTTTCGGTCAACGAGTTTACGTTAATG 673
DB 541 TCAAGAAAGAAAGTTTAGCGGAAGGATATGTTTCGGTCAACGAGTTTACGTTAATG 600

QY 674 AGTAGTGAAGACAAAGCCATCCCTCGGATTTTCATTCGTTGGATGATTGATTAATTTCAAC 733
DB 601 AGTAGTGAAGACAAAGCCATCCCTCGGATTTTCATTCGTTGGATGATTGATTAATTTCAAC 660

QY 734 GTCTCGAAAGTCTACGAGATCGATGCGGAGATCAATCGTGATGCTCTCCAAACCCCAA 793
DB 661 GTCTCGAAAGTCTACGAGATCGATGCGGAGATCAATCGTGATGCTCTCCAAACCCCAA 720

QY 794 AAACCTTTTGACTCTCTCTGTTATGCGCAGGATATATGTAATATCTTAAGTCGT 853
DB 721 AAACCTTTTGACTCTCTCTGTTATGCGCAGGATATATGTAATATCTTAAGTCGT 780

QY 854 TTTTACTTTTTCATCTGTTACTTAATAAAACAAACCCCTTTTTCGGGCAACTTTTCATC 912
DB 781 TTTTACTTTTTCATCTGTTACTTAATAAAACAAACCCCTTTTTCGGGCAACTTTTCATC 839

RESULT 2
AV822206 667 bp mRNA linear EST 01-APR-2002
LOCUS AV822206 RAFL5 Arabidopsis thaliana cDNA clone RAFL05-02-P20 5',
DEFINITION mRNA sequence.

```

```

AV822206
VERSION AV822206.1 GI:19864240
EST.
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 667)
AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
and Shinozaki,K.
TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL Unpublished (2002)
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msekierc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). This clone is in a
modified pBluescript vector as a SstI/XhoI insert. Please visit our
web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for
further details.
FEATURES             Location/Qualifiers
     source            1..667
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="RAFL05-02-P20"
                     /dev_stage="rosette plants"
                     /lab_host="SOLR"
                     /clone_lib="RAFL5"
                     /note="Site 1: SstI; Site 2: XhoI; subjected to
                     dehydration-treated(1,2,5,10,24 hr)"

ORIGIN
Query Match      68.9%; Score 628; DB 1; Length 667;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 628; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 AAGAACCTTACCAAAAGAAACGCTTACGAGACATATCGAGGGAACATCACTTCG 83
DB 25 AAGAACCTTACCAAAAGAAACGCTTACGAGACATATCGAGGGAACATCACTTCG 84

QY 84 TGTAGTTTCAACACGCTTATCATGGAGCTGGATCTGTCTCAAGCTCAAGCCCTCTCTTG 143
DB 85 TGTAGTTTCAACACGCTTATCATGGAGCTGGATCTGTCTCAAGCTCAAGCCCTCTCTTG 144

QY 144 AATCAGCGCGCACCGCGTTACTGTCTGAACTCGCCGCTTCGGGATCGACCCACGAC 203
DB 145 AATCAGCGCGCACCGCGTTACTGTCTGAACTCGCCGCTTCGGGATCGACCCACGAC 204

QY 204 CAATCCAGGCGTGAACACCGTCGACGAATCTCCAAACCGTTGATCGAAACCTTCAAT 263
DB 205 CAATCCAGGCGTGAACACCGTCGACGAATCTCCAAACCGTTGATCGAAACCTTCAAT 264

QY 264 CTCTTCCAGAGAACGAAGGTAATTCGTGTTGGATTCAGCTTCGGAGGCAATCAACATCG 323
DB 265 CTCTTCCAGAGAACGAAGGTAATTCGTGTTGGATTCAGCTTCGGAGGCAATCAACATCG 324

QY 324 CTCTCGCGCGGACATATTTCCGGCGAAGATTAAGTTCTTGTTCTTCAACGCTTCT 383
DB 325 CTCTCGCGCGGACATATTTCCGGCGAAGATTAAGTTCTTGTTCTTCAACGCTTCT 384

QY 384 TGCCCGACACAAACCCACGTCCTTCTACGTTCTGGACAAGTATATGAGATGCTCGGAG 443
DB 385 TGCCCGACACAAACCCACGTCCTTCTACGTTCTGGACAAGTATATGAGATGCTCGGAG 444

QY 444 GTTTCGGAGATTGTGAGTTTTCATCTCATGAAACAAAGAAATGGGACGATGAGTTTATTGA 503

```

```
Db 445 GTTTGGGAGATTCTGAGTTTTCATCTCATGAACAAGAAATGGGACGATGAGTTTATTGA 504
Qy 504 AGATGGGACCAAAATTCATGAAGGCACGCTTTTACCAAAATTTGTCCCATAGAGGATTAGC 563
Db 505 AGATGGGACCAAAATTCATGAAGGCACGCTTTTACCAAAATTTGTCCCATAGAGGATTAGC 564
Qy 564 AGCTGGCAAAATTTGTTCATAGGCAAGGGTCAATTTTTCACAGAGGATCTATCAAGAAAG 623
Db 565 AGCTGGCAAAATTTGTTCATAGGCAAGGGTCAATTTTTCACAGAGGATCTATCAAGAAAG 624
Qy 624 AAAAGTTTACGAGGAGGATGATGGTTC 651
Db 625 AAAAGTTTACGAGGAGGATGATGGTTC 652

RESULT 3
AV785311/c 625 bp mRNA linear EST 28-MAR-2002
LOCUS AV785311 RAPL6 Arabidopsis thaliana cDNA clone RAPL06-14-F12 3',
DEFINITION mRNA sequence.
ACCESSION AV785311 GI:19804101
VERSION AV785311.1
KEYWORDS Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 625)
REFERENCE Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
AUTHORS Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
and Shinozaki,K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
TITLE Unpublished (2002)
JOURNAL Contact: Motoaki Seki
COMMENT RIKEN Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). This clone is in a
modified pBluescript vector as a SstI/XhoI insert. Please visit our
web site (http://www.gsc.riken.go.jp/e/plant/index\_e.html) for
further details.
FEATURES
source Location/Qualifiers
1..625
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAPL06-14-F12"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH10B"
/clone_lib="RAPL6"
/note="Site 1: SstI; Site 2: XhoI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"

ORIGIN
Query Match 68.4%; Score 624; DB 1; Length 625;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 624; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 279 AAGAGGTAATTCGTGGATTTCAGCTTCGGAGGCATCAACATCGCTCTCGCGCGGACA 338
Db 625 AAGAGGTAATTCGTGGATTTCAGCTTCGGAGGCATCAACATCGCTCTCGCGCGGACA 566
Qy 339 TATTTCCGGCGAAGATTAAGGTTCTTGTTCTTCAACGCCCTTTCTTGCCCGACACAACCC 398
|||||
```

```
Db 565 TATTTCCGGCGAAGATTAAGGTTCTTGTTCTTCAACGCCCTTTCTTGCCCGACACAACCC 506
Qy 399 AGTGCCTTTCTCACGTTCTGGCAAGATATATGAGATGCTCGAGGTTTGGGAGATTGTG 458
|||||
Db 505 AGTGCCTTTCTCACGTTCTGGCAAGATATATGAGATGCTCGAGGTTTGGGAGATTGTG 446
Qy 459 AGTTTTTCATCTCATGAAAACAGAAATGGGACGATGAGTTTATTTGAAGATGGGACCAAAAT 518
Db 445 AGTTTTTCATCTCATGAAAACAGAAATGGGACGATGAGTTTATTTGAAGATGGGACCAAAAT 386
Qy 519 TCATGAAGGCACGCTTTTACCAAAATTTGTCATAGAGGATTACGAGCTGGCAAAATGT 578
Db 385 TCATGAAGGCACGCTTTTACCAAAATTTGTCATAGAGGATTACGAGCTGGCAAAATGT 326
Qy 579 TGCATAGGCAAGGGTCAATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTACGAGG 638
Db 325 TGCATAGGCAAGGGTCAATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTACGAGG 266
Qy 639 AAGGATATGGTTTCGTCGCAAGGTTTACGTAATGAGTAGTGAAGCAAAAGCCATCCCT 698
Db 265 AAGGATATGGTTTCGTCGCAAGGTTTACGTAATGAGTAGTGAAGCAAAAGCCATCCCT 206
Qy 699 GCGATTTTCATTCGTTGGATGATTGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATG 758
Db 205 GCGATTTTCATTCGTTGGATGATTGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATG 146
Qy 759 GCGGAGATCACATGCTGATGCTCTCCAAACCCCAAAACTCTTTGACTCTCTCTCTGCTA 818
Db 145 GCGGAGATCACATGCTGATGCTCTCCAAACCCCAAAACTCTTTGACTCTCTCTCTGCTA 86
Qy 819 TTGCCACCGATTATGTAATAATCTTAAGTCCGTTTTTACTTTTTTCTCATCGTTACTAA 878
Db 85 TTGCCACCGATTATGTAATAATCTTAAGTCCGTTTTTACTTTTTTCTCATCGTTACTAA 26
Qy 879 TAAACAACACCCCTTTTTCGCGGC 902
Db 25 TAAACAACACCCCTTTTTCGCGGC 2

RESULT 4
CB258477 641 bp mRNA linear EST 06-NOV-2003
LOCUS CB258477
DEFINITION 05-E012734-014-002-J01-T7R MP12-ADIS-014 Arabidopsis thaliana cDNA
clone MP12p771J012Q 5-PRIME, mRNA sequence.
ACCESSION CB258477
VERSION CB258477.1 GI:32883250
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 641)
REFERENCE Schmid,K.J., Soerensen,T.R., Stracke,R., Torjek,O., Altmann,T.,
AUTHORS Mitchell-Olds,T. and Weisshaar,B.
TITLE Large-scale identification and analysis of genome-wide
single-nucleotide polymorphisms for mapping in Arabidopsis thaliana
JOURNAL Genome Res. 13 (6), 1250-1257 (2003)
MEDLINE 22683290
PUBMED 12799357
COMMENT Contact: Weisshaar B
ADIS DNA core facility at MP1Z
Max-Planck-Institute for Plant Breeding Research
Carl-von-Linne Weg 10, 50829 Koeln, Germany
Fax: 00492215062851
Email: weisshaar@mpiz-koeln.mpg.de
Insert Length: 641 Std Error: 0.00
Plate: 2 row: J column: 01
Seq primer: T7R; CTAATCAGCTCACTATAGGA.
location/Qualifiers
1..641
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
```



```

QY 616 AAGAAAGAAAAGTTTAGCGAGGAAGGATATGTTCCGTCGCAACGAGTTTACGTAATGAG 675
|
Db 297 AAGAAAGAAAAGTTTAGCGAGGAAGGATATGTTCCGTCGCAACGAGTTTACGTAATGAG 238
|
QY 676 TAGTGAAGACAAAGCCATCCCTCGCATTTTCATTCGTTGGATGATTAATTTCAACGT 735
|
Db 237 TAGTGAAGACAAAGCCATCCCTCGCATTTTCATTCGTTGGATGATTAATTTCAACGT 178
|
QY 736 CTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTGATGCTCTCCAAACCCCAAAA 795
|
Db 177 CTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTGATGCTCTCCAAACCCCAAAA 118
|
QY 796 ACTCTTTGACTCTCTCTGCTATGTCACCGGATTAATGTAATTAATCTTAAGTCGGTTT 855
|
Db 117 ACTCTTTGACTCTCTCTGCTATGTCACCGGATTAATGTAATTAATCTTAAGTCGGTTT 58
|
QY 856 TACTTTTTTCTCATCGTTACTATAAACAACACCCCTTTTCCGGCAACTTTTCATC 912
|
Db 57 TACTTTTTTCTCATCGTTACTATAAACAACACCCCTTTTCCGGCAACTTTTCATC 1

RESULT 6
BP613009/c
LOCUS BP613009 RAFL16 Arabidopsis thaliana cDNA clone RAFL16-91-K04 3',
DEFINITION mRNA sequence.
ACCESSION BP613009.1 GI:49264191
VERSION EST.
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 421)
Seki,M., Narusaka,M., Kamiya,A., Ishida,J., Satou,M., Sakurai,T.,
Nakajima,M., Enju,A., Akiyama,K., Oono,Y., Muramatsu,M.,
Hayashizaki,Y., Kawai,J., Carninci,P., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shinagawa,A. and Shinozaki,K.
Functional annotation of a full-length Arabidopsis cDNA collection
Science 296 (5565), 141-145 (2002)
21932900
11910074
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msekirtc.riken.go.jp
reversed clone; please visit our web site
(http://pfweb.gsc.riken.go.jp/) for further details.
FEATURES
source
Location/Qualifiers
1..421
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL16-91-K04"
/lab_host="DH10B"
/clone_lib="RAFL16"
/notes="Site_1: BamHI; Site_2: SalI; dark-grown"

ORIGIN
Query Match 46.2%; Score 421; DB 5; Length 421;
Best Local Similarity 100.0%; Pred. No. 6:3e-222;
Matches 421; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 492 TGAGTTATTGAAGTGGGACCAAAATTCATGAAGCAGCTCTTTACCAAAATTTGCCCA 551
|
Db 421 TGAGTTATTGAAGTGGGACCAAAATTCATGAAGCAGCTCTTTACCAAAATTTGCCCA 362
|
QY 552 TAGAGGATTACGAGCTGGCAAAATTTGTCATAGGCAAGGGTCATTTTTCACAGAGATC 611
|

```

```

Db 361 TAGAGGATTACGAGCTGGCAAAATTTGTCATAGGCAAGGGTCATTTTTCACAGAGATC 302
|
QY 612 TATCAAGAAAGAAAAGTTTAGCGAGGAAGGATATGTTCCGTCGCAACGAGTTTACGTAA 671
|
Db 301 TATCAAGAAAGAAAAGTTTAGCGAGGAAGGATATGTTCCGTCGCAACGAGTTTACGTAA 242
|
QY 672 TGAGTAGTGAAGACAAAGCCATCCCTCGCATTTTCATTCGTTGGATGATTAATTTCA 731
|
Db 241 TGAGTAGTGAAGACAAAGCCATCCCTCGCATTTTCATTCGTTGGATGATTAATTTCA 182
|
QY 732 AGCTCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTGATGCTCTCCAAACCCC 791
|
Db 181 AGCTCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTGATGCTCTCCAAACCCC 122
|
QY 792 AAAAACTCTTTGACTCTCTCTGCTATGTCACCGGATTAATGTAATTAATCTTAAGTCC 851
|
Db 121 AAAAACTCTTTGACTCTCTCTGCTATGTCACCGGATTAATGTAATTAATCTTAAGTCC 62
|
QY 852 GTTTTACTTTTTTCTCATCGTTACTATAAACAACACCCCTTTTCCGGGCAACTTTTCAT 911
|
Db 61 GTTTTACTTTTTTCTCATCGTTACTATAAACAACACCCCTTTTCCGGGCAACTTTTCAT 2

QY 912 C 912
|
Db 1 C 1

RESULT 7
AV822640
LOCUS AV822640 RAFL5 Arabidopsis thaliana cDNA clone RAFL05-09-N03 5',
DEFINITION mRNA sequence.
ACCESSION AV822640
VERSION AV822640.1 GI:19864692
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 572)
Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
and Shinozaki,K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msekirtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). This clone is in a
modified pBluescript vector as a SstI/XhoI insert. Please visit our
web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for
further details.
FEATURES
source
Location/Qualifiers
1..572
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL05-09-N03"
/dev_stage="rosette plants"
/lab_host="SOLR"
/clone_lib="RAFL5"
/notes="Site_1: SstI; Site_2: XhoI; subjected to
dehydration-treated(1,5,10,24 hr)"

ORIGIN
Query Match 43.5%; Score 397; DB 1; Length 572;

```





```
JOURNAL DNA Res. 7 (3), 175-180 (2000)
MEDLINE 20363093
PUBMED 10907847
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1..366
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="R282h03R"
/tissue type="roots"
/clone_lib="Arabidopsis thaliana roots Columbia"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

ORIGIN
Query Match 40.1%; Score 366; DB 1; Length 366;
Best Local Similarity 100.0%; Pred. No. 2.3e-191;
Matches 366; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 AACATCACTTCGTGTAGTTTCAACGCTTATCATGGAGCCTCGATCTGGTACAAGCTCA 131
Db 1 AACATCACTTCGTGTAGTTTCAACGCTTATCATGGAGCCTGGATCTGGTACAGCTCA 60

QY 132 AGCCCTCTCTTGAATCAGCGCGGCCACCGGTTACTCTGTTCGAATCTCGCGCTCCGGGA 191
Db 61 AGCCCTCTCTTGAATCAGCGCGGCCACCGGTTACTCTGTTCGAATCTCGCGCTCCGGGA 120

QY 192 TCGACCCAGCACAATCCAGGCGGTTGAACCGTCCAGCAATCTCCAAACCGTTGATCG 251
Db 121 TCGACCCAGCACAATCCAGGCGGTTGAACCGTCCAGCAATCTCCAAACCGTTGATCG 180

QY 252 AAACCTCTCAAACTCTTCCAGAGAACGAGAGTAATCTGGTTGGATTTCAGCTTCGGAG 311
Db 181 AAACCTCTCAAACTCTTCCAGAGAACGAGAGTAATCTGGTTGGATTTCAGCTTCGGAG 240

QY 312 GCATCAACATCGCTCTCGCGCGCGACATATTTCCGCGGAGAGTAAGGTTCTTGTTCC 371
Db 241 GCATCAACATCGCTCTCGCGCGCGACATATTTCCGCGGAGAGTAAGGTTCTTGTTCC 300

QY 372 TCAACGCTCTTTCGCGCGACACCAACCGTCTTCAAGTCTTGGACAGATATATCG 431
Db 301 TCAACGCTCTTTCGCGCGACACCAACCGTCTTCAAGTCTTGGACAGATATATCG 360

QY 432 AGATGC 437
Db 361 AGATGC 366

RESULT 10
AV820476/c 390 bp mRNA linear EST 01-APR-2002
LOCUS AV820476 RAF111 Arabidopsis thaliana cDNA clone RAF111-10-III 3',
DEFINITION mRNA sequence.
ACCESSION AV820476
VERSION AV820476.1 GI:19862446
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 390)
REFERENCE
AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
Oono,Y., Sakurai,T., Carninci,P., Kawai,S., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
and Shinozaki,K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)

TITLE

JOURNAL Unpublished (2002)
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msekita@tc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified lambdaBamBla-P1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
details.
Location/Qualifiers
1..390
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAF111-10-III"
/dev stage="plants at various developmental stages from
germination to mature seeds"
/lab host="DH10B"
/clone_lib="RAF111"
/note="Site 1: BamHI; Site 2: SalI; subjected to various
treatments (dehydration, cold, high salt, ABA, heat and
UV). Dark-grown plants"

FEATURES
source
1..390
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAF111-10-III"
/dev stage="plants at various developmental stages from
germination to mature seeds"
/lab host="DH10B"
/clone_lib="RAF111"
/note="Site 1: BamHI; Site 2: SalI; subjected to various
treatments (dehydration, cold, high salt, ABA, heat and
UV). Dark-grown plants"

ORIGIN
Query Match 37.4%; Score 341; DB 1; Length 390;
Best Local Similarity 100.0%; Pred. No. 1.8e-177;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 572 AAAATGTTGCATAGCAAGGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTT 631
Db 341 AAAATGTTGCATAGCAAGGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTT 282

QY 632 AGCGAGGAGGATATGTTTCGTCGCAACGAGTTTACGTAATGAGTAGTGAACAAAGCC 691
Db 281 AGCGAGGAGGATATGTTTCGTCGCAACGAGTTTACGTAATGAGTAGTGAACAAAGCC 222

QY 692 ATCCCTCTCGATTTTCATTCGTTGGATGATGATTAATTTCAACGCTCTCGAAAGCTACGAG 751
Db 221 ATCCCTCTCGATTTTCATTCGTTGGATGATGATTAATTTCAACGCTCTCGAAAGCTACGAG 162

QY 752 ATCGATGGCGGAGATCAGATGATGATCTCTCCAAACCCCAAAACTCTTTGACTCTCTC 811
Db 161 ATCGATGGCGGAGATCAGATGATGATCTCTCCAAACCCCAAAACTCTTTGACTCTCTC 102

QY 812 TCTGCTATTGCCACCGATTATATGTAATCTTAAGTCCGTTTACCTTTTCTTCATCG 871
Db 101 TCTGCTATTGCCACCGATTATATGTAATCTTAAGTCCGTTTACCTTTTCTTCATCG 42

QY 872 TTACTATAAAACAAACCCCTTTTTCGGGCAACTTTTCATC 912
Db 41 TTACTATAAAACAAACCCCTTTTTCGGGCAACTTTTCATC 1

RESULT 11
AA404853
LOCUS 28251 Lambda-PRL2 Arabidopsis thaliana cDNA clone 143M11XP 3', mRNA
DEFINITION sequence.
ACCESSION AA404853
VERSION AA404853.1 GI:2062871
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 480)
REFERENCE
```

**AUTHORS** Newman,T., deBruin,J., Green,P., Keestra,K., Kende,H.,  
McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M.,  
Retzel,E. and Somerville,C.  
**TITLE** Genes galore: a summary of methods for accessing results from  
large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
**JOURNAL** Plant Physiol. 106, 1241-1255 (1994)  
**MEDLINE** 95148729  
**PUBMED** 7846151  
**COMMENT** Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PR, Michigan State University, Plant Biology Bldg., E.  
Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22313tcn@bm.cl.msu.edu  
Seq primer: ML3 -21 dye primer.

**FEATURES** Location/Qualifiers  
source  
1..480  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/clone="143M11XP"  
/clone\_lib="Lambda-PRL2"  
/note="Vector: lambda Zip-Lox; Site 1: Sal; Site 2: Not;  
Lambda PRL2 is a cDNA library derived from equal  
quantities of 4 pools of mRNA. The mRNA sources were 1) 7  
day germinated etiolated seedlings; 2) tissue culture  
grown roots; 3) staged plants half with 24 hour light  
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)  
same plants as 3 but aerial tissue (stems, flowers and  
silicles). The vector is BRL's lambda Zip-Lox. The cDNA  
inserts were directionally cloned with Sal-Not arms using  
oligo dT primed cDNA."

## ORIGIN

Query Match 31.7%; Score 289; DB 1; Length 480;  
Best Local Similarity 99.5%; Pred. No. 1.4e-148;  
Matches 389; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 512 CCATAATTCATGAGCAGCTCTTTACCAAAATTTGCCATAGAGGATTACGAGCTGGCA 571  
DB 90 CCATAATTCATGAGCAGCTCTTTACCAAAATTTGCCATAGAGGATTACGAGCTGGCA 149  
  
QY 572 AAAATGTTGCATAGGCAAGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAAGTTT 631  
DB 150 AAAATGTTGCATAGGCAAGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAAGTTT 209  
  
QY 632 AGCGAGGAAGGATATGGTTCCGTGCAACGAGTTTACGTAATGAGTAGTGAAGACAAGCC 691  
DB 210 AGCGAGGAAGGATATGGTTCCGTGCAACGAGTTTACGTAATGAGTAGTGAAGACAAGCC 269  
  
QY 692 ATCCCTCGGATTCATTCGTTGGATGATTGATTAATTTCAACGCTCTCGAAGCTACGAG 751  
DB 270 ATCCCTCGGATTCATTCGTTGGATGATTGATTAATTTCAACGCTCTCGAAGCTACGAG 329  
  
QY 752 ATCGATGGCGGAGATCACATGGTGATGCTCTCCAAACCCCAAAAACCTCTTGACTCTCTC 811  
DB 330 ATCGATGGCGGAGATCACATGGTGATGCTCTCCAAACCCCAAAAACCTCTTGACTCTCTC 389  
  
QY 812 TCTGCTATTGCCACGATTATATGTAATAATCTTTAAGTCGGTTTACTTTTTCTCATCG 871  
DB 390 TCTGCTATTGCCACGATTATATGTAATAATCTTTAAGTCGGTTTACTTTTTCTCATCG 449  
  
QY 872 TTACTAATAAACAACACCCCTTTTCCGGGC 902  
DB 450 TTACTAATAAACAACACCCCTTTTCCGGGC 480

RESULT 12  
AV545429/c  
LOCUS AV545429 335 bp mRNA linear EST 20-FEB-2004

**DEFINITION** AV545429 Arabidopsis thaliana roots Columbia Arabidopsis thaliana  
cDNA clone R282h03F 3', mRNA sequence.  
**ACCESSION** AV545429  
**VERSION** AV545429.1 GI:8716843  
**KEYWORDS** EST.  
**SOURCE** Arabidopsis thaliana (thale cress)  
**ORGANISM** Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
**REFERENCE** 1 (bases 1 to 335)  
**AUTHORS** Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.  
**TITLE** A large scale analysis of cDNA in Arabidopsis thaliana: Generation  
of 12,028 non-redundant expressed sequence tags from normalized and  
size-selected cDNA libraries  
**JOURNAL** DNA Res. 7 (3), 175-180 (2000)  
**MEDLINE** 20363093  
**PUBMED** 10907847  
**COMMENT** Contact: Erika Asamizu  
The First Laboratory for Plant Gene Research  
Kazusa DNA Research Institute  
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

**FEATURES** Location/Qualifiers  
source  
1..335  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/clone="R282h03F"  
/tissue\_type="roots"  
/clone\_lib="Arabidopsis thaliana roots Columbia"  
/note="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2:  
XhoI"

## ORIGIN

Query Match 30.6%; Score 279; DB 1; Length 335;  
Best Local Similarity 100.0%; Pred. No. 5e-143;  
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 634 CGAGGAAGGATATGGTTCCGTGCAACGAGTTTACGTAATGAGTAGTGAAGACAAGCCAT 693  
DB 314 CGAGGAAGGATATGGTTCCGTGCAACGAGTTTACGTAATGAGTAGTGAAGACAAGCCAT 255  
  
QY 694 CCCCTGCCATTTCAATTCGTTGGATGATTGATAATTTCAACGCTCTCGAAGCTACGAGAT 753  
DB 254 CCCCTGCCATTTCAATTCGTTGGATGATTGATAATTTCAACGCTCTCGAAGCTACGAGAT 195  
  
QY 754 CGATGGCGGAGATCACATGGTGATGCTCTCCAAACCCCAAAAACCTCTTGACTCTCTCTC 813  
DB 194 CGATGGCGGAGATCACATGGTGATGCTCTCCAAACCCCAAAAACCTCTTGACTCTCTCTC 135  
  
QY 814 TGCTATTGCCACCGATTATATGTAATAATCTTAAAGTCGGTTTACTTTTTCTCATCGTT 873  
DB 134 TGCTATTGCCACCGATTATATGTAATAATCTTAAAGTCGGTTTACTTTTTCTCATCGTT 75  
  
QY 874 ACTAATAAACAACACCCCTTTTCCGGCAACTTTTCATC 912  
DB 74 ACTAATAAACAACACCCCTTTTCCGGCAACTTTTCATC 36

RESULT 13  
BP646852/c  
LOCUS BP646852 382 bp mRNA linear EST 27-JUN-2004  
**DEFINITION** BP646852 RAFL19 Arabidopsis thaliana cDNA clone RAFL19-73-109 3',  
mRNA sequence.  
**ACCESSION** BP646852  
**VERSION** BP646852.1 GI:49298322  
**KEYWORDS** EST.  
**SOURCE** Arabidopsis thaliana (thale cress)  
**ORGANISM** Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 382)  
Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T., Nakajima, M., Enju, A., Akiyama, K., Oono, Y., Muramatsu, M., Hayaishizaki, Y., Kawai, J., Carninci, P., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A. and Shinozaki, K.  
Functional annotation of a full-length Arabidopsis cDNA collection  
Science 296 (5565), 141-145 (2002)  
21932900  
MEDLINE  
PUBMED

REFERENCE  
AUTHORSTITLE  
JOURNALMEDLINE  
PUBMED

## COMMENT

Contact: Motoaki Seki  
Plant Functional Genomics Research Group  
RIKEN Genomic Sciences Center  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-4359  
Fax: 81-298-36-9060  
Email: mseki@rtc.riken.go.jp  
reversed clone; Please visit our web site  
(http://pfweb.gsc.riken.go.jp/) for further details.  
Location/Qualifiers

## FEATURES

source

1. .382  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone="RAFL19-73-I09"  
/tissue type="mixture of silique and flower"  
/lab\_host="DH10B"  
/clone\_lib="RAFL19"  
/notes="Site\_1: BamHI; Site\_2: SalI; Subtraction Library"

## ORIGIN

Query Match 28.8%; Score 263; DB 5; Length 382;  
Best Local Similarity 99.7%; Pred. No. 3.9e-134;  
Matches 313; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 587 CAAGGTCATTTTTCACAGAGATCTATCAAGAAAGAAAAGTTTAGCGAGGAGATAT 646  
Db 344 CAAGGTCATTTTTCACAGAGATCTATCAAGAAAGAAAAGTTTAGCGAGGAGATAT 285  
  
QY 647 GGTTCGGTCCACGAGTTTACGTAAAGTAGTAGTGAAGCAAGCCATCCCTGCCATTTC 706  
Db 284 GGTTCGGTCCACGAGTTTACGTAAAGTAGTAGTGAAGCAAGCCATCCCTGCCATTTC 225  
  
QY 707 ATTGCTTGGATGATTCATTAATTTCAACGCTCTGAAAGTCTAGGAGATCGATGCGGAGAT 766  
Db 224 ATTGCTTGGATGATTCATTAATTTCAACGCTCTGAAAGTCTAGGAGATCGATGCGGAGAT 165  
  
QY 767 CACATGGTGATGCTCTCCAAACCCCAAAACTTTTGACTCTCTCTGCTATTGCCACC 826  
Db 164 CACATGGTGATGCTCTCCAAACCCCAAAACTTTTGACTCTCTCTGCTATTGCCACC 105  
  
QY 827 GATTATATGTAATATCTTAAGTCGCTTTTACTTTTCTCATCGTTACTATAAACA 886  
Db 104 GATTATATGTAATATCTTAAGTCGCTTTTACTTTTCTCATCGTTACTATAAACA 45  
  
QY 887 ACCCTTTTTCGGG 900  
Db 44 ACCCTTTTTCGGG 31

RESULT 14  
AI993246LOCUS  
DEFINITIONACCESSION  
AI993246VERSION  
AI993246.1KEYWORDS  
EST.SOURCE  
Arabidopsis thalianaORGANISM  
Arabidopsis thaliana

AI993246 361 bp mRNA linear EST 08-SEP-1999  
701495839 A. thaliana, Ohio State clone set Arabidopsis thaliana  
cDNA clone 701495839, mRNA sequence.  
AI993246.1 GI:5840151  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE  
AUTHORSTITLE  
JOURNAL

## COMMENT

1 (bases 1 to 361)  
Chen, J., Hillyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D., Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P., Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D., Tan, R., Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C., Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobrega, A., Murry, L., Turner, C., Krikorian, S., Elder, L. and Hanson, D.  
Arabidopsis thaliana Gene Expression MicroArray  
Unpublished (1999)  
Contact: David Smoller, Ph.D.  
Genome Systems, Inc., a wholly owned subsidiary of Incyte  
Pharmaceuticals, Inc.  
4633 World Parkway Circle, St. Louis, MO 63134, USA  
Tel: 877-577-2733  
Fax: 314-427-3324  
Email: service@genomesystems.com.  
Location/Qualifiers

## FEATURES

source

1. .361  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone="701495839"  
/clone\_lib="A. thaliana, Ohio State clone set"  
/note="cDNA library was made from selected clones from the  
Arabidopsis thaliana Ohio State clone set."

## ORIGIN

Query Match 28.4%; Score 259; DB 1; Length 361;  
Best Local Similarity 99.4%; Pred. No. 6.5e-132;  
Matches 359; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 14 CAACCAAGCAAGAACCTTTAAACCAAAAGAAAACGCTTACGAGACATATGGAGAGAAA 73  
Db 1 CAACCAAGCAAGAACCTTTAAACCAAAAGAAAACGCTTACGAGACATATGGAGAGAAA 60  
  
QY 74 CATCATCTCGTGTAGTTTCAACAGCTTATCATGGAGCTGATCTGGTACAGCTCAAG 133  
Db 61 CATCATCTCGTGTAGTTTCAACAGCTTATCATGGAGCTGATCTGGTACAGCTCAAG 120  
  
QY 134 CCCCTCTTGAATCAGCCGCGCACCGCTTACTGCTGCGAACTCGCCGCTCCGGGATC 193  
Db 121 CCCCTCTTGAATCAGCCGCGCACCGCTTACTGCTGCGAACTCGCCGCTCCGGGATC 180  
  
QY 194 GACCCACCAACCAATCCAGCGCTTGAAACCGTCGACGAATACTCCAAACCGTTGATCGAA 253  
Db 181 GACCCACCAACCAATCCAGCGCTTGAAACCGTCGACGAATACTCCAAACCGTTGATCGAA 240  
  
QY 254 ACCCTCAAATCTTTCCAGAGAACGAGAGTAATTTCTGGTTGGATTTCAGCTTCGAGGC 313  
Db 241 ACCCTCAAATCTTTCCAGAGAACGAGAGTAATTTCTGGTTGGATTTCAGCTTCGAGGC 300  
  
QY 314 ATCAACATCGCTCTCGCCGCGCACATATTTCCGGGGAAGATTAGGTTCTTGTGTCTC 373  
Db 301 ATCAACATCGCTCTCGCCGCGCACATATTTCCGGGGAAGATTAGGTTCTTGTGTCTC 360  
  
QY 374 A 374  
Db 361 A 361

RESULT 15  
AV559909/cLOCUS  
DEFINITIONACCESSION  
AV559909VERSION  
AV559909.1KEYWORDS  
EST.SOURCE  
Arabidopsis thalianaORGANISM  
Arabidopsis thaliana

AV559909 308 bp mRNA linear EST 23-FEB-2004  
AV559909 Arabidopsis thaliana green silique Columbia Arabidopsis  
thaliana cDNA clone SQ125f05f 3', mRNA sequence.  
AV559909  
AV559909.1 GI:8731335  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

|  |   |
|--|---|
| rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. |   |
| REFERENCE  | 1 (bases 1 to 308)  |
| AUTHORS  | Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.  |
| TITLE  | A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries   |
| JOURNAL  | DNA Res. 7 (3), 175-180 (2000)  |
| MEDLINE  | 20363093  |
| PUBMED   | 10907847  |
| COMMENT  | Contact: Erika Asamizu<br>The First Laboratory for Plant Gene Research<br>Kazusa DNA Research Institute<br>Yana 1532-3, Kisarazu, Chiba 292-0812, Japan<br>Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/. |
| FEATURES   | Location/Qualifiers   |
| source   | 1..308  |
|  | /organism="Arabidopsis thaliana"  |
|  | /mol_type="mRNA"  |
|  | /ecotype="Columbia"   |
|  | /db_xref="taxon:3702"   |
|  | /clones="SQ125f05f"   |
|  | /tissue_type="green siliques"   |
|  | /clone_lib="Arabidopsis thaliana green siliques Columbia"   |
|  | /note=Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"   |
| ORIGIN   |   |
| Query Match  | 27.1%; Score 247; DB 1; Length 308;   |
| Best Local Similarity  | 100.0%; Pred. No. 3e-125;   |
| Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |   |
| Qy   | 666 ACGTAATGAGTAGAGACAAAGCCCATCCCTCGATTTCATTCGTTGGATGATTCATA 725  |
| Db   | 308 ACGTAATGAGTAGAGACAAAGCCCATCCCTCGATTTCATTCGTTGGATGATTCATA 249  |
| Qy   | 726 ATTTCAACGTCGAAAAGTCATCAGAGATCGATGCGGAGATCACATGGTGATGCTCTCCA 785   |
| Db   | 248 ATTTCAACGTCGAAAAGTCATCAGAGATCGATGCGGAGATCACATGGTGATGCTCTCCA 189   |
| Qy   | 786 AACCCCAAAACTCTTTGACTCTCTCTGCTATTTGCCACGATTATATGTAATAATCTT 845   |
| Db   | 188 AACCCCAAAACTCTTTGACTCTCTCTGCTATTTGCCACGATTATATGTAATAATCTT 129   |
| Qy   | 846 AAGTCGGTTTTACTTTTTTCTCATCGTTACTAATAAAAAACAAACCCCTTTTTCGGGCAAC 905   |
| Db   | 128 AAGTCGGTTTTACTTTTTTCTCATCGTTACTAATAAAAAACAAACCCCTTTTTCGGGCAAC 69  |
| Qy   | 906 TTTCATC 912   |
| Db   | 68 TTTCATC 62   |

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 06:32:02 ; Search time 499.287 Seconds  
(without alignments)  
10813.032 Million cell updates/sec

Title: US-10-780-002-36

Perfect score: 912

Sequence: 1 atcagctattctcaacaa.....ttttcgggcaactttcatc 912

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Match | Length | DB ID        | Description         |
|------------|-------|-------|--------|--------------|---------------------|
| 1          | 912   | 100.0 | 994    | 8 ADA26534   | Ada26534 Arabidops  |
| 2          | 861   | 94.4  | 1056   | 3 AAC47895   | Aac47895 Arabidops  |
| 3          | 844   | 92.5  | 909    | 3 AAC47800   | Aac47800 Arabidops  |
| 4          | 432   | 47.4  | 483    | 10 ADE82083  | Ades2083 Arabidops  |
| 5          | 20    | 2.2   | 8592   | 6 ABL33982   | Ab133982 Human imm  |
| 6          | 19    | 2.1   | 371    | 4 AAI80037   | Aai80037 Human pol  |
| 7          | 19    | 2.1   | 387    | 10 ADD49350  | Add49350 Human lun  |
| 8          | 19    | 2.1   | 393    | 12 ADL03170  | Adl03170 DNA encod  |
| 9          | 19    | 2.1   | 432    | 13 ADR64551  | Adr64551 Cotton cd  |
| 10         | 19    | 2.1   | 440    | 10 ADD49311  | Add49311 Human lun  |
| 11         | 19    | 2.1   | 624    | 6 ABK39793   | Abk39793 DNA encod  |
| 12         | 19    | 2.1   | 624    | 8 ACA12122   | Aca12122 Human His  |
| 13         | 19    | 2.1   | 624    | 8 ACA03308   | Aca03308 Lung canc  |
| 14         | 19    | 2.1   | 624    | 10 ADH47433  | Adh47433 Human lun  |
| 15         | 19    | 2.1   | 624    | 13 ADJ211352 | Adj211352 Human lun |
| 16         | 19    | 2.1   | 636    | 6 ABK39629   | Abk39629 cDNA enco  |
| 17         | 19    | 2.1   | 636    | 8 ACAL1958   | Acal1958 Human lun  |
| 18         | 19    | 2.1   | 636    | 8 ACA03144   | Aca03144 Lung canc  |
| 19         | 19    | 2.1   | 636    | 10 ADH47187  | Adh47187 Human lun  |
| 20         | 19    | 2.1   | 636    | 13 ADJ211106 | Adj211106 Human lun |

|      |    |     |       |             |                     |
|------|----|-----|-------|-------------|---------------------|
| C 21 | 19 | 2.1 | 1310  | 6 ABQ54338  | Abq54338 Human ova  |
| C 22 | 19 | 2.1 | 1633  | 5 AAF68398  | Aaf68398 Human lun  |
| C 23 | 19 | 2.1 | 1633  | 5 AAF68850  | Aaf68850 Human lun  |
| C 24 | 19 | 2.1 | 1633  | 6 ABK38761  | Abk38761 cDNA enco  |
| C 25 | 19 | 2.1 | 1633  | 6 ABK38309  | Abk38309 cDNA enco  |
| C 26 | 19 | 2.1 | 1633  | 8 ACAL0638  | Acal0638 Human lun  |
| C 27 | 19 | 2.1 | 1633  | 8 ACAL1090  | Acal1090 Human lun  |
| C 28 | 19 | 2.1 | 1633  | 8 ACA02276  | Aca02276 Lung canc  |
| C 29 | 19 | 2.1 | 1633  | 8 ABX99589  | Abx99589 Lung canc  |
| C 30 | 19 | 2.1 | 1633  | 10 ADH45835 | Adh45835 Human lun  |
| C 31 | 19 | 2.1 | 1633  | 10 ADH46307 | Adh46307 Human lun  |
| C 32 | 19 | 2.1 | 1633  | 12 ADE72372 | Ades72372 Human lun |
| C 33 | 19 | 2.1 | 1633  | 13 ADR25282 | Adr25282 Breast ca  |
| C 34 | 19 | 2.1 | 1633  | 13 ADJ19754 | Adj19754 Human lun  |
| C 35 | 19 | 2.1 | 1633  | 13 ADJ20226 | Adj20226 Human lun  |
| C 36 | 19 | 2.1 | 2179  | 5 AAS87251  | Aas87251 DNA encod  |
| C 37 | 19 | 2.1 | 3266  | 4 AAH34750  | Aah34750 Human col  |
| C 38 | 19 | 2.1 | 3300  | 4 ABL03694  | Ab103694 Drosophil  |
| C 39 | 19 | 2.1 | 3521  | 8 ABX76282  | Abx76282 Lung canc  |
| C 40 | 19 | 2.1 | 3565  | 12 ADN41807 | Adn41807 Novel hum  |
| C 41 | 19 | 2.1 | 9325  | 13 ADR84244 | Adr84244 Aspergill  |
| C 42 | 19 | 2.1 | 32186 | 5 AAS34422  | Aas34422 Human DNA  |
| C 43 | 19 | 2.1 | 38855 | 12 ADN41808 | Adn41808 Novel hum  |
| C 44 | 19 | 2.1 | 63248 | 11 ACN45100 | Acn45100 Mouse gen  |
| C 45 | 18 | 2.0 | 283   | 6 ABL74493  | Ab174493 Corn tags  |

#### ALIGNMENTS

##### RESULT 1

ADA26534

ID ADA26534 standard; DNA; 984 BP.

XX ADA26534;

XX 20-NOV-2003 (first entry)

XX Arabidopsis thaliana AtSBL5 gene homologous to tobacco SABP2 gene.  
XX salicylic acid-binding protein; SABP2; disease resistance; plant;  
XX modulator; tobacco; ds; gene.

XX Arabidopsis thaliana.

XX WO2003016551-A2.

XX 27-FEB-2003.

XX 16-AUG-2002; 2002WO-US026312.

XX 16-AUG-2001; 2001US-0312863P.

XX (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

XX Klessig DF, Kumar D;

XX WPI; 2003-278578/27.

XX New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for  
XX producing transgenic plants with increased resistance to disease, or for  
XX screening SABP2 modulators that confer enhanced resistance of plants to  
XX disease.  
XX Claim 65; Page; 98pp; English.  
XX The invention relates to a novel isolated 1079 bp salicylic acid-binding  
XX protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid  
XX protein, sequences that specifically hybridize with the nucleic acid, the  
XX complement of the nucleic acid or a natural allelic variant of the  
XX sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for  
XX conferring increased resistance to disease in plants, or for screening  
XX modulators of SABP2, which confer increased or enhanced resistance of

CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also  
CC useful for producing transgenic plants with increased resistance to  
CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful  
CC for studying the molecular mechanisms responsible for the modulation of  
CC salicylic acid-mediated disease resistance in plants. This sequence  
CC represents the gene encoding the Arabidopsis thaliana SABP2 AtSB2L5 which  
CC has homology to the gene (ADA26498) encoding the tobacco SABP2 protein.  
CC (Note: this sequence is not given in the specification but is derived  
CC from the Genbank accession number given by the inventors).  
XX

SQ Sequence 984 BP; 291 A; 230 C; 200 G; 263 T; 0 U; 0 Other;

Query Match 100.0%; Score 912; DB 8; Length 984;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 912; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |     |  |     |
|----|-----|--|-----|
| Qy | 1   | ATCAGCGTATTCTCAAAACAGCAAGAACCTTAAACCAAAAGAAACCGTCTACGGAGACA    | 60  |
| Db | 1   | ATCAGCGTATTCTCAAAACAGCAAGAACCTTAAACCAAAAGAAACCGTCTACGGAGACA    | 60  |
| Qy | 61  | TATGAGAGGAAACATCACTTCGTGTAGTTTCAACAAGCTTATCATGAGAGCCTGGATCTG   | 120 |
| Db | 61  | TATGAGAGGAAACATCACTTCGTGTAGTTTCAACAAGCTTATCATGAGAGCCTGGATCTG   | 120 |
| Qy | 121 | GTACAAGCTCAAGCCCTCTTGAATCAGCGGCGCACCGCTTACTGTCTGCGAACTCGC      | 180 |
| Db | 121 | GTACAAGCTCAAGCCCTCTTGAATCAGCGGCGCACCGCTTACTGTCTGCGAACTCGC      | 180 |
| Qy | 181 | CGCCTCCGGGATCGACCCACGACCAATCCAGCGCGTTGAAACCGTCGACGAATATCCAA    | 240 |
| Db | 181 | CGCCTCCGGGATCGACCCACGACCAATCCAGCGCGTTGAAACCGTCGACGAATATCCAA    | 240 |
| Qy | 241 | ACCGTTGATCGAAACCCCTCAAAATCTCTTCAGAGAACGAGGTAATCTCTGGTTGGATT    | 300 |
| Db | 241 | ACCGTTGATCGAAACCCCTCAAAATCTCTTCAGAGAACGAGGTAATCTCTGGTTGGATT    | 300 |
| Qy | 301 | CAGCTTCGGAGGATCAATCGCTCTCGCGCGACATATTTCCGGCGGAGATTAAAGT        | 360 |
| Db | 301 | CAGCTTCGGAGGATCAATCGCTCTCGCGCGACATATTTCCGGCGGAGATTAAAGT        | 360 |
| Qy | 361 | TCTTGTGTTCTCAACGCCCTTCTGCCCGACACCAACCCACGTCCTTCTCAGGTTCTGGA    | 420 |
| Db | 361 | TCTTGTGTTCTCAACGCCCTTCTGCCCGACACCAACCCACGTCCTTCTCAGGTTCTGGA    | 420 |
| Qy | 421 | CAAGTATATGAGAGATCGCTCGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAACAAG  | 480 |
| Db | 421 | CAAGTATATGAGAGATCGCTCGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAACAAG  | 480 |
| Qy | 481 | AAATGGGACGATGAGTTTATTCAGATGGGACCAAAATTCATGAAGGCAACGCTTTTACCA   | 540 |
| Db | 481 | AAATGGGACGATGAGTTTATTCAGATGGGACCAAAATTCATGAAGGCAACGCTTTTACCA   | 540 |
| Qy | 541 | AAATTTGTCCTATAGAGGATTACGAGTTCGCAAAATTTGTCATAGGCAAGGTCATTTT     | 600 |
| Db | 541 | AAATTTGTCCTATAGAGGATTACGAGTTCGCAAAATTTGTCATAGGCAAGGTCATTTT     | 600 |
| Qy | 601 | CACAGAGGATCTATCAAAAGAAAGAAAGTTTACGAGGAAGGATATGTTTCGGTCAACG     | 660 |
| Db | 601 | CACAGAGGATCTATCAAAAGAAAGAAAGTTTACGAGGAAGGATATGTTTCGGTCAACG     | 660 |
| Qy | 661 | AGTTTACGTAAGTAGTAGTAGAGCAAAAGCCATCCCTCGGATTTTCATTCGTTGGATGAT   | 720 |
| Db | 661 | AGTTTACGTAAGTAGTAGTAGAGCAAAAGCCATCCCTCGGATTTTCATTCGTTGGATGAT   | 720 |
| Qy | 721 | TGATTAATTTCAACGCTCTCGAAAGTCTTACGAGATCGATGGCGGAGATCACATGGTGATGT | 780 |
| Db | 721 | TGATTAATTTCAACGCTCTCGAAAGTCTTACGAGATCGATGGCGGAGATCACATGGTGATGT | 780 |
| Qy | 781 | CTCCAAACCCCAAAACTCTTTGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT  | 840 |
| Db | 781 | CTCCAAACCCCAAAACTCTTTGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT  | 840 |
| Qy | 841 | ATCTTAAGTCCGTTTACTTTTTTCTCATCGTTACTTAATAAAACCAACCCCTTTTCCGG    | 900 |

|          |   |   |     |
|----------|---|---|-----|
| Db       | 841   | ATCTTAAGTCCGTTTACTTTTTTCTCATCGTTACTTAATAAAACCAACCCCTTTTCCGG | 900 |
| Qy       | 901   | GCAACTTTTCATC 912   |     |
| Db       | 901   | GCAACTTTTCATC 912   |     |
| RESULT 2 |   |   |     |
| ID       | AAC47895  | standard; DNA; 1056 BP.                                     |     |
| AC       | AAC47895;   |   |     |
| XX       |   |   |     |
| DT       | 18-OCT-2000   | (first entry)   |     |
| XX       |   |   |     |
| DE       | Arabidopsis thaliana  | DNA fragment SEQ ID NO: 55510.                              |     |
| KW       | Hybridisation assay; genetic mapping; gene expression control;          |   |     |
| KW       | protein identification; signal transduction pathway; metabolic pathway; |   |     |
| XX       | promoter; termination sequence; ss.                                     |   |     |
| OS       | Arabidopsis thaliana.   |   |     |
| XX       |   |   |     |
| PN       | EP1033405-A2.   |   |     |
| XX       |   |   |     |
| PD       | 06-SEP-2000.  |   |     |
| XX       |   |   |     |
| PF       | 25-FEB-2000; 2000EP-00301439.   |   |     |
| XX       |   |   |     |
| PR       | 25-FEB-1999; 99US-0121825P.   |   |     |
| PR       | 05-MAR-1999; 99US-0123180P.   |   |     |
| PR       | 09-MAR-1999; 99US-0123548P.   |   |     |
| PR       | 23-MAR-1999; 99US-0125788P.   |   |     |
| PR       | 25-MAR-1999; 99US-0126264P.   |   |     |
| PR       | 29-MAR-1999; 99US-0126785P.   |   |     |
| PR       | 01-APR-1999; 99US-0127462P.   |   |     |
| PR       | 06-APR-1999; 99US-0128234P.   |   |     |
| PR       | 08-APR-1999; 99US-0128714P.   |   |     |
| PR       | 16-APR-1999; 99US-0129845P.   |   |     |
| PR       | 19-APR-1999; 99US-0130077P.   |   |     |
| PR       | 21-APR-1999; 99US-0130449P.   |   |     |
| PR       | 23-APR-1999; 99US-0130510P.   |   |     |
| PR       | 23-APR-1999; 99US-0130891P.   |   |     |
| PR       | 28-APR-1999; 99US-0131449P.   |   |     |
| PR       | 30-APR-1999; 99US-0132048P.   |   |     |
| PR       | 30-APR-1999; 99US-0132407P.   |   |     |
| PR       | 04-MAY-1999; 99US-0132484P.   |   |     |
| PR       | 05-MAY-1999; 99US-0132485P.   |   |     |
| PR       | 06-MAY-1999; 99US-0132486P.   |   |     |
| PR       | 06-MAY-1999; 99US-0132487P.   |   |     |
| PR       | 07-MAY-1999; 99US-0132863P.   |   |     |
| PR       | 11-MAY-1999; 99US-0134256P.   |   |     |
| PR       | 14-MAY-1999; 99US-0134218P.   |   |     |
| PR       | 14-MAY-1999; 99US-0134219P.   |   |     |
| PR       | 14-MAY-1999; 99US-0134221P.   |   |     |
| PR       | 14-MAY-1999; 99US-0134370P.   |   |     |
| PR       | 18-MAY-1999; 99US-0134768P.   |   |     |
| PR       | 19-MAY-1999; 99US-0134941P.   |   |     |
| PR       | 20-MAY-1999; 99US-0135124P.   |   |     |
| PR       | 21-MAY-1999; 99US-0135353P.   |   |     |
| PR       | 24-MAY-1999; 99US-0135629P.   |   |     |
| PR       | 25-MAY-1999; 99US-0136021P.   |   |     |
| PR       | 27-MAY-1999; 99US-0136392P.   |   |     |
| PR       | 28-MAY-1999; 99US-0136782P.   |   |     |
| PR       | 01-JUN-1999; 99US-0137222P.   |   |     |
| PR       | 03-JUN-1999; 99US-0137528P.   |   |     |
| PR       | 04-JUN-1999; 99US-0137502P.   |   |     |
| PR       | 07-JUN-1999; 99US-0137724P.   |   |     |
| PR       | 08-JUN-1999; 99US-0138094P.   |   |     |
| PR       | 10-JUN-1999; 99US-0138540P.   |   |     |
| PR       | 10-JUN-1999; 99US-0138847P.   |   |     |
| PR       | 14-JUN-1999; 99US-0139119P.   |   |     |

PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142820P.  
PR 12-JUL-1999; 99US-0142877P.  
PR 13-JUL-1999; 99US-0143342P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.

PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 94.4%; Score 861; DB 3; Length 1056;  
Best Local Similarity 99.9%; Pred. No. 0;

Matches 911; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCAGCTATTCTCAACAGCAAGAACCTTAACCAAAAAGAAACGTTACCGAGACA 60

Db 128 ATCAGCTATTCTCAACAGCAAGAACCTTAACCAAAAAGAAACGTTACCGAGACA 187

QY 61 TATGAGAGGAACATCACTTCGTGTTAGTTTCAACCGCTTATCATGAGCGCTGATCTG 120  
DB |||||  
QY 188 TATGAGAGGAACATCACTTCGTGTTAGTTTCAACCGCTTATCATGAGCGCTGATCTG 247  
DB |||||  
QY 121 GTACAAGCTCAAGCCCTCCTTGAATCAGCCGGCCACCGGTTACTGCTGCGAACTCGC 180  
DB |||||  
QY 248 GTACAAGCTCAAGCCCTCCTTGAATCAGCCGGCCACCGGTTACTGCTGCGAACTCGC 307  
DB |||||  
QY 181 CGCTCCGGGATCGACCCACGACCAATCAGCGCGCTTGAACCGTCGACGAATATCCAA 240  
DB |||||  
QY 308 CGCTCCGGGATCGACCCACGACCAATCAGCGCGCTTGAACCGTCGACGAATATCCAA 367  
DB |||||  
QY 241 ACCGTTGATCGAAACCGCTCAATCTCTCCAGAGAACGAGGTAATCTCGTTGGATT 300  
DB |||||  
QY 368 ACCGTTGATCGAAACCGCTCAATCTCTCCAGAGAACGAGGTAATCTCGTTGGATT 427  
DB |||||  
QY 301 CAGCTTCGGAGGCATCAACATCGCTCTCGCCCGCACATATTTCCGGCGAAGATTAAGGT 360  
DB |||||  
QY 428 CAGCTTCGGAGGCATCAACATCGCTCTCGCCCGCACATATTTCCGGCGAAGATTAAGGT 487  
DB |||||  
QY 361 TCTTGTGTTCTCAACCGCTTCTGCCGACACAAACCGCGCTTCTCAGCTTCGGA 420  
DB |||||  
QY 488 TCTTGTGTTCTCAACCGCTTCTGCCGACACAAACCGCGCTTCTCAGCTTCGGA 547  
DB |||||  
QY 421 CAGTATATGAGATCGCTGGAGGTTTGGAGATTGTGAGTTTTCATCTCATGAACAAG 480  
DB |||||  
QY 548 CAGTATATGAGATCGCTGGAGGTTTGGAGATTGTGAGTTTTCATCTCATGAACAAG 607  
DB |||||  
QY 481 AAATGGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGGCACGCTTTACCA 540  
DB |||||  
QY 608 AAATGGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGGCACGCTTTACCA 667  
DB |||||  
QY 541 AAATGTCCCATAGAGATTACGAGCTGGCAAAATGTTGCATAGGCAAGGTCATTTT 600  
DB |||||  
QY 668 AAATGTCCCATAGAGATTACGAGCTGGCAAAATGTTGCATAGGCAAGGTCATTTT 727  
DB |||||  
QY 601 CACAGGATCTATCAAGAGAAAGTTTAGCGGGAAGGATATGTTGGTGCAACG 660  
DB |||||  
QY 728 CACAGAGGATCTATCAAGAGAAAGTTTAGCGGGAAGGATATGTTGGTGCAACG 787  
DB |||||  
QY 661 AGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTCGCATTTTCATTTCGTTGGATGAT 720  
DB |||||  
QY 788 AGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTCGCATTTTCATTTCGTTGGATGAT 847  
DB |||||  
QY 721 TCATAATTTCAACGCTCGAAAGTCTACGAGATCGATGCGGAGATCACATGGTATGCT 780  
DB |||||  
QY 848 TGATAATTTCAACGCTCGAAAGTCTACGAGATCGATGCGGAGATCACATGGTATGCT 907  
DB |||||  
QY 781 CTCCAAACCCAAAACCTTTTGACTCTCTCTGCTATTGCCCGATTATATGTAATA 840  
DB |||||  
QY 908 CTCCAAACCCAAAACCTTTTGACTCTCTCTGCTATTGCCCGATTATATGTAATA 967  
DB |||||  
QY 841 ATCTTAAGTCCGTTTACTTTTCTCATCGTTACTTAATAAAACAAACCCCTTTTCCGG 900  
DB |||||  
QY 968 ATCTTAAGTCCGTTTACTTTTCTCATCGTTACTTAATAAAACAAACCCCTTTTCCGG 1027  
DB |||||  
QY 901 GCAACTTTCATC 912  
DB |||||  
QY 1028 GCAACTTTCATC 1039  
DB |||||

RESULT 3  
AAC47800  
ID AAC47800 standard; DNA; 909 BP.  
XX  
AC AAC47800;  
XX  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 55168.  
XX  
XX Hybridization assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW

KW promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
OS  
XX EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126284P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.



|    |              |  |     |              |                |
|----|--------------|--|-----|--------------|----------------|
| PR | 28-JUN-1999; | 99US-0140823P.   | PR  | 10-SEP-1999; | 99US-0153070P. |
| PR | 29-JUN-1999; | 99US-0140991P.   | PR  | 13-SEP-1999; | 99US-0153758P. |
| PR | 30-JUN-1999; | 99US-0141287P.   | PR  | 15-SEP-1999; | 99US-0154018P. |
| PR | 01-JUL-1999; | 99US-0141842P.   | PR  | 16-SEP-1999; | 99US-0154039P. |
| PR | 01-JUL-1999; | 99US-0142154P.   | PR  | 20-SEP-1999; | 99US-0154779P. |
| PR | 02-JUL-1999; | 99US-0142055P.   | PR  | 22-SEP-1999; | 99US-0155139P. |
| PR | 06-JUL-1999; | 99US-0142390P.   | PR  | 23-SEP-1999; | 99US-0155486P. |
| PR | 08-JUL-1999; | 99US-0142803P.   | PR  | 24-SEP-1999; | 99US-0155659P. |
| PR | 09-JUL-1999; | 99US-0142920P.   | PR  | 28-SEP-1999; | 99US-0156458P. |
| PR | 12-JUL-1999; | 99US-0142977P.   | PR  | 29-SEP-1999; | 99US-0156596P. |
| PR | 13-JUL-1999; | 99US-0143542P.   | PR  | 04-OCT-1999; | 99US-0157117P. |
| PR | 14-JUL-1999; | 99US-0143624P.   | PR  | 05-OCT-1999; | 99US-0157553P. |
| PR | 15-JUL-1999; | 99US-0144005P.   | PR  | 06-OCT-1999; | 99US-0157865P. |
| PR | 16-JUL-1999; | 99US-0144086P.   | PR  | 07-OCT-1999; | 99US-0158029P. |
| PR | 16-JUL-1999; | 99US-0144325P.   | PR  | 08-OCT-1999; | 99US-0158232P. |
| PR | 19-JUL-1999; | 99US-0144331P.   | PR  | 12-OCT-1999; | 99US-0158369P. |
| PR | 19-JUL-1999; | 99US-0144332P.   | PR  | 13-OCT-1999; | 99US-0159293P. |
| PR | 19-JUL-1999; | 99US-0144333P.   | PR  | 13-OCT-1999; | 99US-0159294P. |
| PR | 19-JUL-1999; | 99US-0144334P.   | PR  | 14-OCT-1999; | 99US-0159295P. |
| PR | 19-JUL-1999; | 99US-0144335P.   | PR  | 14-OCT-1999; | 99US-0159329P. |
| PR | 20-JUL-1999; | 99US-0144352P.   | PR  | 14-OCT-1999; | 99US-0159330P. |
| PR | 20-JUL-1999; | 99US-0144352P.   | PR  | 14-OCT-1999; | 99US-0159331P. |
| PR | 20-JUL-1999; | 99US-0144632P.   | PR  | 14-OCT-1999; | 99US-0159637P. |
| PR | 21-JUL-1999; | 99US-0144884P.   | PR  | 14-OCT-1999; | 99US-0159638P. |
| PR | 21-JUL-1999; | 99US-0145086P.   | PR  | 18-OCT-1999; | 99US-0159584P. |
| PR | 21-JUL-1999; | 99US-0145088P.   | PR  | 21-OCT-1999; | 99US-0160741P. |
| PR | 22-JUL-1999; | 99US-0145085P.   | PR  | 21-OCT-1999; | 99US-0160767P. |
| PR | 22-JUL-1999; | 99US-0145087P.   | PR  | 21-OCT-1999; | 99US-0160768P. |
| PR | 22-JUL-1999; | 99US-0145089P.   | PR  | 21-OCT-1999; | 99US-0160770P. |
| PR | 22-JUL-1999; | 99US-0145192P.   | PR  | 21-OCT-1999; | 99US-0160814P. |
| PR | 23-JUL-1999; | 99US-0145145P.   | PR  | 22-OCT-1999; | 99US-0160980P. |
| PR | 23-JUL-1999; | 99US-0145218P.   | PR  | 22-OCT-1999; | 99US-0160981P. |
| PR | 23-JUL-1999; | 99US-0145224P.   | PR  | 22-OCT-1999; | 99US-0160985P. |
| PR | 26-JUL-1999; | 99US-0145276P.   | PR  | 25-OCT-1999; | 99US-0161404P. |
| PR | 27-JUL-1999; | 99US-0145913P.   | PR  | 25-OCT-1999; | 99US-0161405P. |
| PR | 27-JUL-1999; | 99US-0145918P.   | PR  | 25-OCT-1999; | 99US-0161406P. |
| PR | 27-JUL-1999; | 99US-0145919P.   | PR  | 26-OCT-1999; | 99US-0161359P. |
| PR | 28-JUL-1999; | 99US-0145951P.   | PR  | 26-OCT-1999; | 99US-0161360P. |
| PR | 02-AUG-1999; | 99US-0146386P.   | PR  | 26-OCT-1999; | 99US-0161361P. |
| PR | 02-AUG-1999; | 99US-0146388P.   | PR  | 28-OCT-1999; | 99US-0161920P. |
| PR | 02-AUG-1999; | 99US-0146389P.   | PR  | 28-OCT-1999; | 99US-0161992P. |
| PR | 03-AUG-1999; | 99US-0147038P.   | PR  | 28-OCT-1999; | 99US-0161993P. |
| PR | 04-AUG-1999; | 99US-0147204P.   | PR  | 29-OCT-1999; | 99US-0162142P. |
| PR | 04-AUG-1999; | 99US-0147302P.   |     |              |                |
| PR | 05-AUG-1999; | 99US-0147192P.   |     |              |                |
| PR | 05-AUG-1999; | 99US-0147260P.   |     |              |                |
| PR | 06-AUG-1999; | 99US-0147303P.   |     |              |                |
| PR | 06-AUG-1999; | 99US-0147416P.   |     |              |                |
| PR | 09-AUG-1999; | 99US-0147493P.   |     |              |                |
| PR | 09-AUG-1999; | 99US-0147935P.   |     |              |                |
| PR | 10-AUG-1999; | 99US-0148171P.   |     |              |                |
| PR | 11-AUG-1999; | 99US-0148319P.   |     |              |                |
| PR | 12-AUG-1999; | 99US-0148341P.   |     |              |                |
| PR | 13-AUG-1999; | 99US-0148565P.   |     |              |                |
| PR | 13-AUG-1999; | 99US-0148684P.   |     |              |                |
| PR | 16-AUG-1999; | 99US-0149368P.   |     |              |                |
| PR | 17-AUG-1999; | 99US-0149175P.   |     |              |                |
| PR | 18-AUG-1999; | 99US-0149426P.   |     |              |                |
| PR | 20-AUG-1999; | 99US-0149722P.   |     |              |                |
| PR | 20-AUG-1999; | 99US-0149723P.   |     |              |                |
| PR | 20-AUG-1999; | 99US-0149829P.   |     |              |                |
| PR | 23-AUG-1999; | 99US-0149902P.   |     |              |                |
| PR | 23-AUG-1999; | 99US-0149930P.   |     |              |                |
| PR | 25-AUG-1999; | 99US-0150566P.   |     |              |                |
| PR | 26-AUG-1999; | 99US-0150884P.   |     |              |                |
| PR | 27-AUG-1999; | 99US-0151065P.   |     |              |                |
| PR | 27-AUG-1999; | 99US-0151066P.   |     |              |                |
| PR | 27-AUG-1999; | 99US-0151080P.   |     |              |                |
| PR | 30-AUG-1999; | 99US-0151303P.   |     |              |                |
| PR | 31-AUG-1999; | 99US-0151438P.   |     |              |                |
| PR | 01-SEP-1999; | 99US-0151930P.   |     |              |                |
| PR | 07-SEP-1999; | 99US-0152363P.   |     |              |                |
| Qy | 4            | AGGCTATTCTCAACAAGCAAGAACCTTTAACCAAAAAGAAACCTCTACGGAGACATAT   | 63  |              |                |
| Db | 10           | ACGCTATTCTCAACAAGCAAGAACCTTTAACCAAAAAGAAACCTCTACGGAGACATAT   | 69  |              |                |
| Qy | 64           | GGAGAGGAAACATCACTTCGTGTTAGTTCACAACGCTTATCATGAGCCTGGATCTGGTA  | 123 |              |                |
| Db | 70           | GGAGAGGAAACATCACTTCGTGTTAGTTCACAACGCTTATCATGAGCCTGGATCTGGTA  | 129 |              |                |
| Qy | 124          | CAAGCTCAAGCCCTCTTGTGAATCAGCGGCGCACCGCGTTACTGTGTGAACTCGCCGC   | 183 |              |                |
| Db | 130          | CAAGCTCAAGCCCTCTTGTGAATCAGCGGCGCACCGCGTTACTGTGTGAACTCGCCGC   | 189 |              |                |
| Qy | 184          | CTCCGGGATCGACCCACGACCAATCCAGCCGTTGAACCGCTCGACGAATCTCCAAACC   | 243 |              |                |
| Db | 190          | CTCCGGGATCGACCCACGACCAATCCAGCCGTTGAACCGCTCGACGAATCTCCAAACC   | 249 |              |                |
| Qy | 244          | GTTCATCGAAACCCCTCAAAATCTTCCAGAGAACGAAGAGGTAATCTGGTTGGATTTCAG | 303 |              |                |
| Db | 250          | GTTCATCGAAACCCCTCAAAATCTTCCAGAGAACGAAGAGGTAATCTGGTTGGATTTCAG | 309 |              |                |
| Qy | 304          | CTTCGGAGGCATCAACATCGCTCTCGCCGCGACATATTTCCGGCGAAGATTAAAGTTCT  | 363 |              |                |
| Db | 310          | CTTCGGAGGCATCAACATCGCTCTCGCCGCGACATATTTCCGGCGAAGATTAAAGTTCT  | 369 |              |                |

Query Match 92.5%; Score 844; DB 3; Length 909;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 844; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 364 TGTGTTCTCAACGGCTTCTTGCCCGACACCAACCCACGTCGCTTCTCAGCTTCTGGACAA 423  
DB |||||  
QY 370 TGTGTTCTCAACGGCTTCTTGCCCGACACCAACCCACGTCGCTTCTCAGCTTCTGGACAA 429  
DB |||||  
QY 424 GTATATGAGATGCGCTGGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAACAAGAAA 483  
DB |||||  
QY 430 GTATATGAGATGCGCTGGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAACAAGAAA 489  
DB |||||  
QY 484 TGGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGGCACGTCCTTTTACCAAAA 543  
DB |||||  
QY 490 TGGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGGCACGTCCTTTTACCAAAA 549  
DB |||||  
QY 544 TTGTCCTCATAGAGATTACGACTGGCAAAATGTTGATAGGCAAGGGTCATTTTTCAC 603  
DB |||||  
QY 604 AGAGGATCTATCAAGAAGAAAAGTTTAGCGAGGAAGATATGTTGCGTGCAACGAGT 663  
DB |||||  
QY 610 AGAGGATCTATCAAGAAGAAAAGTTTAGCGAGGAAGATATGTTGCGTGCAACGAGT 669  
DB |||||  
QY 664 TTACGTAATGAGTAGTGAAGACAAAGCCATCCCTGCGATTTCATTCGTTGGATGATGA 723  
DB |||||  
QY 670 TTAAGTAATGAGTAGTGAAGACAAAGCCATCCCTGCGATTTCATTCGTTGGATGATGA 729  
DB |||||  
QY 724 TAAATTTCAACGTCGAAAAGTCTACGAGATCGATGCGGAGATCACATGCTGATGCTCTC 783  
DB |||||  
QY 730 TAAATTTCAACGTCGAAAAGTCTACGAGATCGATGCGGAGATCACATGCTGATGCTCTC 789  
DB |||||  
QY 784 CAAACCCCAAAACCTTTTGACTCTCTCTGCTATTTGCGCCGATTATATGTAATAATC 843  
DB |||||  
QY 790 CAAACCCCAAAACCTTTTGACTCTCTCTGCTATTTGCCACCGATTATATGTAATAATC 849  
DB |||||  
QY 844 TTAAT 847  
DB |||||  
QY 850 TTAAT 853  
DB |||||  
RESULT 4  
ADE82083  
ID ADE82083 standard; cdNA; 483 BP.  
XX AC ADE82083;  
XX DT 29-JAN-2004 (first entry)  
XX DE Arabidopsis thaliana expressed polynucleotide seq id 854.  
XX KW genetically modified organism; transgenic organism; plant;  
KW inhibitor testing; activator testing; modifier testing; fungicide;  
KW insecticide; genetic function; genetic regulation; cellular metabolism;  
KW gene; ss.  
XX OS Arabidopsis thaliana.  
XX PN US2003115639-A1.  
XX PD 19-JUN-2003.  
XX PP 26-JAN-2001; 2001US-00770961.  
XX PR 27-JAN-2000; 2000US-0178466P.  
XX (GORL/) GORLACH J.  
PA (ANY/) AN Y.  
PA (HAMI/) HAMILTON C M.  
PA (PRIC/) PRICE J L.  
PA (RAIN/) RAINES T M.  
PA (YUYU/) YU Y.  
PA (RAME/) RAMEAKA J G.  
PA (PAGE/) PAGE A.  
PA (MATH/) MATHEN A V.  
PA (LEDF/) LEDFORD B L.  
PA (WOES/) WOESSNER J P.

PA (HAAS/) HAAS W D.  
PA (GARC/) GARCIA C A.  
PA (KRIC/) KRICKER M.  
PA (SLAT/) SLATER T.  
PA (DAVI/) DAVIS K R.  
PA (ALLE/) ALLEN K.  
PA (HOFF/) HOFFMAN N.  
PA (HURB/) HURBAN P.  
XX  
Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y,  
PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;  
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;  
PI Hurban P;  
XX WPI; 2003-810930/76.  
XX Novel Arabidopsis thaliana nucleic acids useful for generating  
PT genetically modified transgenic organisms, for screening biologically  
PT active agents such as fungicides, insecticides.  
XX Claim 1; SEQ ID NO 854; 44pp; English.  
XX The invention describes a nucleic acid (I) comprising a sequence capable  
CC of hybridising under stringent conditions to any one of 999 fully defined  
CC Arabidopsis thaliana sequences (I) as given in specification e.g., 360,  
CC 1137, 455, 219, 472 nucleotides, etc, or its fragment. (I) is useful as a  
CC hybridisation probe to complementary molecules in a cDNA library. (I) is  
CC also useful for generating genetically modified and transgenic organisms,  
CC usually plant cells and plants. A protein encoded by (I) is useful in  
CC screening assays to determine the effect of candidate inhibitors,  
CC activators or modifiers of the gene product. The protein is also useful  
CC for screening biologically active agents e.g., fungicides and  
CC insecticides. A genetically modified cell, comprising an exogenous  
CC nucleic acid, where the nucleic acid comprises transcription regulatory  
CC sequences operably linked to a sequence capable of hybridising under  
CC stringent conditions to (I) is useful in the study of genetic function  
CC and regulation, for alteration of the cellular metabolism and for  
CC screening compounds that may affect the biological function of the gene  
CC or gene product. This sequence represents an Arabidopsis thaliana  
CC polynucleotide of the invention.  
XX SQ Sequence 483 BP; 129 A; 127 C; 112 G; 115 T; 0 U; 0 Other;  
Query Match 47.4%; Score 432; DB 10; Length 483;  
Best Local Similarity 99.8%; Pred. No. 4.3e-214;  
Matches 482; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 26 GAACCTTAACCAAAAGAAAACGCTACGGAGACATATGGAGAGAAACATCATTCTG 85  
DB 1 GAACCTTAACCAAAAGAAAACGCTACGGAGACATATGGAGAGAAACATCATTCTG 60  
QY 86 TTAGTTCAACACGCTTATCATGGAGCCTGGATCTGGTCAAGCTCAAGCCCTCTTGA 145  
DB 61 TTAGTTCAACACGCTTATCATGGAGCCTGGATCTGGTCAAGCTCAAGCCCTCTTGA 120  
QY 146 TCAGCCGGCCACCGGTTACTCTCGAACTCCGCGCTCCGGATCGACCCAGCA 205  
DB 121 TCAGCCGGCCACCGGTTACTCTCGAACTCCGCGCTCCGGATCGACCCAGCA 180  
QY 206 ATCCAGGCGGTTGAAAACCGTCGCAATACCTCCAAACCGTTGATCGAAACCTCAAA 265  
DB 181 ATCCAGGCGGTTGAAAACCGTCGCAATACCTCCAAACCGTTGATCGAAACCTCAAA 240  
QY 266 CTTCCAGAGAACGAAGAGGTAATCTGTTGGATTACGTTCCGAGGATCAACATCGT 325  
DB 241 CTTTCAGAGAACGAAGAGGTAATCTGTTGGATTACGTTCCGAGGATCAACATCGT 300  
QY 326 CTCGCGCGCGACATATTTCCGCGAGAGATTAAAGTTCTTGTGTTCTCAAGCCCTCTT 385  
DB 301 CTCGCGCGCGACATATTTCCGCGAGAGATTAAAGTTCTTGTGTTCTCAAGCCCTCTT 360  
QY 386 CCGGACACAAACCCAGCTGCTCTTCTCAGCTTCTGGCAAGATATATGGAGATGCTCGAG 445  
DB |||||

```

Db 361 CCGACACACACCACGCTCTCTCACGTTCTGGACAAGTATATGAGATGCGCTGGAGGT 420
QY 446 TTGGGAGATTGTGAGTTTTCATCTCATGAACAAGAAATGGGACGATGAGTTTATTGGAAG 505
Db 421 TTGGGAGATTGTGAGTTTTCATCTCATGAACAAGAAATGGGACGATGAGTTTATTGGAAG 480
QY 506 ATG 508
Db 481 ATG 483

RESULT 5
ABL33982/c
ID ABL33982 standard; DNA; 8592 BP.
XX
AC ABL33982;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 1955.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antitartaroclerotic; anianaemic; cytosatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP007537.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (SPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
XX
PS Claim 1; SEQ ID NO 1955; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
XX
SQ Sequence 8592 BP; 2473 A; 132 C; 1657 G; 4330 T; 0 U; 0 Other;

Query Match 2.2%; Score 20; DB 6; Length 8592;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 781 CTCCTAACCCCAAACTCT 800
Db 2474 CTCCTAACCCCAAACTCT 2455

```

```

RESULT 6
AAI80037/c
ID AAI80037 standard; cDNA; 371 BP.
XX
AC AAI80037;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 97.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
XX
PN P-PSDB; AAO00106.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
PS Claim 1; SEQ ID NO 97; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 371 BP; 104 A; 43 C; 55 G; 169 T; 0 U; 0 Other;

Query Match 2.1%; Score 19; DB 4; Length 371;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 465 CATCTCATGAACAAGAAA 483
Db 275 CATCTCATGAACAAGAAA 257

RESULT 7
ADD49350/c
ID ADD49350 standard; cDNA; 387 BP.
XX
AC ADD49350;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human lung cancer associated cDNA 61483100.
XX

```



KW pathogen resistance; pest resistance; yield; photosynthesis; seed oil;  
 KW stress resistance.  
 XX Gossypium hirsutum.  
 XX US2004181830-A1.  
 PN 16-SEP-2004.  
 XX 29-JAN-2004; 2004US-00767795.  
 XX 07-MAY-2001; 2001US-00849529.  
 PR 12-DEC-2001; 2001US-00021323.  
 XX (KOVA/) KOVALIC D K.  
 PA (ZHOU/) ZHOU Y.  
 PA (CAO/) CAO Y.  
 XX Kovalic DK, Zhou Y, Cao Y;  
 PI WPI; 2004-667718/65.  
 DR New recombinant nucleic acid molecules and polypeptides from Gossypium  
 PT hirsutum, useful for producing plants with improved biological  
 PT characteristics (e.g. improved plant cold or drought tolerance).  
 XX Claim 1; SEQ ID NO 5332; 14pp; English.  
 PS The invention relates to a recombinant polynucleotide comprising any of  
 XX the 58798 Cotton plant cDNA sequences mentioned in the specification.  
 CC Also a recombinant polypeptide comprising any of the 58798 amino acid  
 CC sequences mentioned in the specification and producing a plant having an  
 CC improved property. Producing a plant having an improved property  
 CC comprises transforming a plant with a recombinant construct comprising a  
 CC promoter region functional in a plant cell operably joined to a  
 CC polynucleotide comprising a coding sequence for a polypeptide associated  
 CC with the property, and growing the transformed plant. The polypeptide is  
 CC useful for improving plant cold tolerance, manipulating growth rate in  
 CC plant cells by modification of the cell cycle pathway, improving plant  
 CC drought tolerance, providing increased resistance to plant disease,  
 CC producing galactanmannan (or lignin or plant growth regulators), improving  
 CC plant heat tolerance, improving plant tolerance to herbicides, increasing  
 CC the rate of homologous recombination in plants, improving plant tolerance  
 CC to extreme osmotic conditions or to pathogens or pests, improving yield  
 CC by modification of photosynthesis, modifying seed oil or protein yield  
 CC and/or content, improving yield by modification of carbohydrate, nitrogen  
 CC or phosphorus use and/or uptake, or improving yield by providing improved  
 CC plant growth and development under at least one stress condition. The  
 CC polynucleotide and polypeptide may also be used in recombinant DNA  
 CC constructs, in physical arrays of molecules, as plant breeding markers,  
 CC or in computer-based storage and analysis systems. The present sequence  
 CC is a Cotton plant cDNA of the invention. NOTE: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequences.html?DocID=20040181830. However only 6585  
 CC polynucleotide sequences were available, the remaining 52113  
 CC polynucleotides and all 58798 protein sequences were not present.  
 XX Sequence 432 BP; 127 A; 81 C; 79 G; 145 T; 0 U; 0 Other;  
 SQ Query Match 2.1%; Score 19; DB 13; Length 432;  
 Best Local Similarity 100.0%; Pred. No. 48;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 880 AAAACAACCCCTTTTCC 898  
 DB 1 AAAACAACCCCTTTTCC 19  
 RESULT 10  
 ADD49911/c  
 ID ADD49911 standard; cDNA; 440 BP.  
 XX

AC ADD49911;  
 XX 15-JAN-2004 (first entry)  
 XX Human lung cancer associated cDNA 61824027.  
 DE Human; ss; lung cancer antigen; cytostatic; lung cancer; gene therapy;  
 KW vaccine; T-cell; tumour.  
 XX Homo sapiens.  
 XX US2003194764-A1.  
 PN 16-OCT-2003.  
 XX 04-APR-2002; 2002US-00116712.  
 PF 05-APR-2001; 2001US-0282289P.  
 PR 05-OCT-2001; 2001US-0327511P.  
 XX (CORI-) CORIXA CORP.  
 XX Bangur CS, Switzer A;  
 PI WPI; 2003-844452/78.  
 DR New isolated polypeptides and polynucleotides useful for diagnosing,  
 XX preventing and treating cancer, particularly lung cancer.  
 PT Claim 1; SEQ ID NO 643; 250pp; English.  
 PS The invention relates to an isolated polynucleotide (a) comprising any of  
 XX the 666 fully defined nucleotide sequences appearing as ADD49269 -  
 CC ADD49935, ADD49936 and ADD49938, complements of (a); sequences of at  
 CC least 20 contiguous residues of (a); sequences that hybridise to (a)  
 CC under highly stringent conditions; sequences having at least 75 or 90%  
 CC identity to (a); or degenerate variants of (a). Also included are an  
 CC isolated polypeptide (b) (comprising: sequences encoded by the new  
 CC polynucleotide; any of the 4 amino acid sequences fully defined in the  
 CC specification; or sequences having at least 70 or 90% identity to the  
 CC sequence in (a) or (b)), an expression vector comprising the above  
 CC polynucleotide operably linked to an expression control sequence, a host  
 CC cell transformed or transfected with the above expression vector, an  
 CC isolated antibody, or its antigen-binding fragment, that specifically  
 CC binds to the above polypeptide, an oligonucleotide that hybridises to the  
 CC above-mentioned nucleotide sequences under highly stringent conditions, a  
 CC fusion protein comprising at least one polypeptide cited above, detecting  
 CC the presence of a cancer in a patient (comprising: obtaining a biological  
 CC sample from the patient; contacting the biological sample with a binding  
 CC agent that binds to the polypeptide, or with the oligonucleotide cited  
 CC above; detecting in the sample an amount of the polypeptide that binds to  
 CC the binding agent, or an amount of a polynucleotide that hybridises to  
 CC the oligonucleotide; and comparing the amount of polypeptide, or  
 CC polynucleotide that hybridises to the oligonucleotide, to a predetermined  
 CC cut-off value and then determining the presence of a cancer in the  
 CC patient), a method for stimulating and/or expanding T-cells specific for  
 CC a tumour protein (comprising contacting T-cells with the above  
 CC polypeptide, polynucleotide or antigen-presenting cells that express the  
 CC polypeptide, under conditions and for a time sufficient to permit the  
 CC stimulation and/or expansion of T-cells), an isolated T-cell population  
 CC comprising T-cells prepared by the method, a composition comprising a  
 CC first component selected from physiological carriers and  
 CC immunostimulants, and a second component selected from the above  
 CC polypeptide, polynucleotide, antibody, fusion protein, T-cell population  
 CC and antigen-presenting cells that express the above polypeptide,  
 CC stimulating an immune response in a patient (comprising administering to  
 CC the patient the above composition) treating lung cancer in a patient  
 CC (comprising administering to the patient the above composition and a  
 CC diagnostic kit (comprising: at least one oligonucleotide cited above; or  
 CC at least one antibody cited above and a detection reagent, where the  
 CC detection reagent comprises a reporter group). The composition and  
 CC methods are useful in diagnosing, preventing and treating cancer,  
 CC particularly lung cancer. The present sequence is a lung cancer -

```
CC associated antigen cDNA of the invention.
XX
SQ Sequence 440 BP; 144 A; 61 C; 122 G; 113 T; 0 U; 0 Other;
    Query Match          2.1%; Score 19; DB 10; Length 440;
    Best Local Similarity 100.0%; Pred. No. 48;
    Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGGCAACTTT 908
Db 137 CCTTTTCCGGGCAACTTT 119

RESULT 11
ABK39793/c
ID ABK39793 standard; cDNA; 624 BP.
XX
AC ABK39793;
XX
DT 21-MAY-2002 (first entry)
XX
DE DNA encoding recombinant protein L548S.
XX
KW Lung tumour; cancer; T cell; immune response stimulator; cytostatic;
KW Gene; ss.
XX
OS Homo sapiens.
XX
PN WO200204514-A2.
XX
PD 17-JAN-2002.
XX
PF 10-JUL-2001; 2001WO-US022058.
XX
PR 11-JUL-2000; 2000US-00614124.
PR 29-AUG-2000; 2000US-00651563.
PR 08-SEP-2000; 2000US-00658824.
PR 26-SEP-2000; 2000US-00671325.
PR 06-OCT-2000; 2000US-00677419.
PR 30-OCT-2000; 2000US-00702705.
PR 13-DEC-2000; 2000US-00736457.
PR 03-MAY-2001; 2001US-00849626.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;
PI Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS, Mcnabb A;
PI Wang A, Fanger N, Switzer A, Mcneill PD, Clapper JD;
XX
WPI; 2002-164634/31.
DR P-PSDB; AAU85606.
XX
XX Novel polynucleotide encoding a lung tumor polypeptide useful for
PT stimulating and/or expanding T cells specific for a tumor protein.
XX
XX Claim 9; SEQ ID NO 1914; 223pp; English.
XX
XX The invention describes an isolated polynucleotide and polypeptide useful
CC for stimulating and/or expanding T cells specific for a tumour protein
CC for determining the presence of a cancer in a patient. A composition
CC containing the polynucleotide and/or polypeptide is useful for treating a
CC lung cancer in a patient. The polypeptide is useful for removing tumour
CC cells from a biological sample. The polynucleotide is also useful as
CC probe or primer to detect the level of mRNA encoding a tumour protein.
CC This sequence encodes a lung tumour associated protein or protein
CC fragment, described in the method of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 624 BP; 221 A; 110 C; 178 G; 115 T; 0 U; 0 Other;
    Query Match          2.1%; Score 19; DB 6; Length 624;

Best Local Similarity 100.0%; Pred. No. 48;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTCCGGGCAACTTT 908
Db 559 CCTTTTCCGGGCAACTTT 541

RESULT 12
ACA12122/c
ID ACA12122 standard; cDNA; 624 BP.
XX
AC ACA12122;
XX
DT 06-JUN-2003 (first entry)
XX
DE Human His tagged lung cancer associated protein L548S, cDNA.
XX
KW Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
KW T cell expansion; CD4; CD8; RAI2; gene.
XX
OS Homo sapiens.
XX
PN US2002197669-A1.
XX
PD 26-DEC-2002.
XX
PF 03-MAY-2001; 2001US-00849626.
XX
PR 13-DEC-2000; 2000US-00736457.
XX
PA (BANG/) BANGUR C S.
PA (FANG/) FANGER G R.
PA (WANG/) WANG A.
PA (WANG/) WANG T.
PA (SWIT/) SWITZER A P.
PA (MCNE/) MCNEILL P D.
PA (CLAP/) CLAPPER J D.
XX
PI Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, Mcneill PD;
PI Clapper JD;
XX
WPI; 2003-352750/33.
DR P-PSDB; ABU69588.
XX
XX Novel lung cancer polynucleotide encoding lung cancer protein, useful for
PT detecting the presence of lung cancer in a patient, and in pharmaceutical
XX compositions, e.g. vaccines, for treating lung cancer.
XX
XX Example 14; Page; 72pp; English.
XX
XX The invention relates to a polynucleotide encoding a lung tumour protein,
CC comprising a sequence selected from any of the 14 sequences mentioned in
CC the specification, or a sequence (S2) mentioned in specification,
CC complement of S1, sequences consisting of at least 20 contiguous residues
CC of S1, sequences that hybridise to S1, sequences having 75%, preferably
CC 90%, identity to S1, or degenerate variants of S1. Also included are an
CC isolated polypeptide (comprising a sequence (S3) selected from any one of
CC the 4 amino acid sequences mentioned in the specification, a sequence
CC encoded by the polynucleotide, or sequences having at least 70%,
CC preferably 90%, identity to a sequence encoded by the polynucleotide), an
CC expression vector comprising the polynucleotide operably linked to an
CC expression control sequence, a host cell transformed or transfected with
CC the vector, an isolated antibody (or its antigen-binding fragment) that
CC specifically binds to the polypeptide, detecting the presence of a cancer
CC in a patient, a fusion protein comprising the polypeptide, an
CC oligonucleotide that hybridises to S1 under moderately stringent
CC conditions, stimulating and/or expanding T cells specific for a tumour
CC protein (comprising contacting T cells with the polynucleotide, protein
CC or antigen-presenting cells, under conditions and for a time sufficient
CC to permit the stimulation and/or expansion of T cells) and inhibiting the
CC development of a cancer in a patient (by incubating CD4+ and/or CD8+ T
CC cells isolated from a patient with the polynucleotide, protein or antigen
```

CC presenting cells that express the polynucleotide, such that T cells  
 CC proliferate, administering to the patient an effective amount of the  
 CC proliferated T cells, and thus inhibiting the development of a cancer in  
 CC the patient. The polynucleotide, protein and cells are useful in a  
 CC composition for stimulating an immune response in a patient, and for  
 CC treating a cancer in a patient (particularly lung cancer). The  
 CC oligonucleotide is useful for determining the presence of a cancer in a  
 CC patient. The protein and oligonucleotides are useful in pharmaceutical  
 CC compositions, e.g. vaccines. The polynucleotide is also useful as a probe  
 CC or primer for nucleic acid hybridisation, and in the design and  
 CC preparation of ribozyme molecules for inhibiting expression of tumour  
 CC polypeptides and proteins in tumour cells. An amplified portion of the  
 CC polynucleotide is useful for isolating a full-length gene from a suitable  
 CC library. The present sequence encodes a fusion protein of human RA12 with  
 CC the protein product of a cDNA (full length, extended or partial) isolated  
 CC from a library derived from lung tumour/cancer cells. Note: The sequence  
 CC data for this patent did not form part of the printed specification, but  
 CC was obtained in electronic format directly from the USPTO at  
 CC seqdata.uspto.gov/sequence.html?docId=20020197669  
 XX  
 SQ Sequence 624 BP; 221 A; 110 C; 178 G; 115 T; 0 U; 0 Other;

Query Match 2.1%; Score 19; DB 8; Length 624;  
 Best Local Similarity 100.0%; Pred. No. 48;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTTCGGGCAACTTT 908  
 DB 559 CCTTTTTCGGGCAACTTT 541

RESULT 13  
 ACA03308/c  
 ID ACA03308 standard; DNA; 624 BP.

XX ACA03308;

XX 22-MAY-2003 (first entry)

XX Lung cancer therapy and diagnosis associated DNA #15.

XX Lung cancer; cytostatic; vaccine; gene therapy; cancer; gene; ds.

XX Homo sapiens.

XX US2002172952-A1.

XX 21-NOV-2002.

XX 10-JUL-2001; 2001US-00902941.

XX 30-JUN-1999; 99US-00346492.

XX 15-OCT-1999; 99US-00419356.

XX 17-DEC-1999; 99US-00466867.

XX 30-DEC-1999; 99US-00476300.

XX 06-MAR-2000; 2000US-00519642.

XX 22-MAR-2000; 2000US-00533077.

XX 10-APR-2000; 2000US-00546259.

XX 27-APR-2000; 2000US-00560406.

XX 05-JUN-2000; 2000US-00589184.

XX 11-JUL-2000; 2000US-00614124.

XX 29-AUG-2000; 2000US-00651563.

XX 08-SEP-2000; 2000US-00658824.

XX 26-SEP-2000; 2000US-00671325.

XX 06-OCT-2000; 2000US-00677419.

XX 30-OCT-2000; 2000US-00702705.

XX 13-DEC-2000; 2000US-00736457.

XX 03-MAY-2001; 2001US-00849626.

XX (CORI-) CORIXA CORP.

XX Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW;

PI Durham M, Carter D, Fanger GR, Vedvick TS, Bangur CS, McNabb A;

XX WPI; 2003-328427/31.

XX New polynucleotide, useful for preparing a composition for treating or  
 PT inhibiting development of cancer, e.g. lung cancer.

XX Example 14; SEQ ID NO 1914; 82pp; English.

XX The invention describes an isolated polynucleotide comprising one of 32  
 CC sequences' complement or degenerate variants of them. The polynucleotide  
 CC is useful for preparing a composition e.g. a vaccine or for gene therapy,  
 CC for treating or inhibiting development of cancer, e.g. lung cancer. This  
 CC sequence represents a polynucleotide associated with the compositions and  
 CC methods for the therapy and diagnosis of lung cancer

SQ Sequence 624 BP; 221 A; 110 C; 178 G; 115 T; 0 U; 0 Other;

Query Match 2.1%; Score 19; DB 8; Length 624;  
 Best Local Similarity 100.0%; Pred. No. 48;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 890 CCTTTTTCGGGCAACTTT 908

DB 559 CCTTTTTCGGGCAACTTT 541

RESULT 14

ADH47433/c

ID ADH47433 standard; DNA; 624 BP.

XX AC ADH47433;

XX 25-MAR-2004 (first entry)

XX Human lung tumour DNA, SEQ ID No 1914.

XX lung cancer; tumour; cytostatic; immunostimulant; vaccine; gene therapy;  
 KW human; ds.

XX Homo sapiens.

XX WO2003037267-A2.

XX 08-MAY-2003.

XX 28-OCT-2002; 2002WO-US034777.

XX 29-OCT-2001; 2001US-00017754.

XX 28-MAR-2002; 2002US-00113872.

XX (CORI-) CORIXA CORP.

XX Henderson RA, Wang T, Watanabe Y, Kalos MD, Sleath PR;

PI Johnson JC, Retter MW, Durham M, Carter D, Fanger GR, Vedvick TS;

PI Bangur CS, McNabb A;

XX WPI; 2003-468346/44.

XX New polypeptides and encoding polynucleotides, useful for diagnosing,  
 PT preventing and/or treating lung cancer.

XX Example 14; SEQ ID NO 1914; 258pp; English.

XX The invention relates to novel compositions and methods for the therapy  
 CC and diagnosis of cancer, particularly lung cancer. The compositions  
 CC comprise one or more lung tumour polypeptides, immunogenic portions  
 CC thereof, polynucleotides that encode such polypeptides, antigen presenting  
 CC cells that express such polypeptides, and T cells that are specific for  
 CC cells expressing such polypeptides. The novel compositions have  
 CC cytostatic and immunostimulant activity. The lung tumour antigens can be  
 CC used in the creation of a vaccine. The polynucleotides that encode the  
 CC lung tumour polypeptides can be used in gene therapy to help in the  
 CC treatment of lung tumours. This polynucleotide represents a human lung

CC tumour DNA sequence of the invention. This sequence was not shown in the  
CC specification. It has been taken from a World Intellectual Property  
CC Organization CD ROM supplied with the specification.

XX  
SQ Sequence 624 BP; 221 A; 110 C; 178 G; 115 T; 0 U; 0 Other;  
Query Match 2.1%; Score 19; DB 10; Length 624;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 890 CCTTTTTCGGGCAACTTT 908  
|||||  
Db 559 CCTTTTTCGGGCAACTTT 541

RESULT 15  
ADJ21352/c  
ID ADJ21352 standard; DNA; 624 BP.

XX AC ADJ21352;  
XX  
DT 04-NOV-2004 (first entry)  
XX Human lung cancer-related L548S coding region DNA.  
DE  
XX lung tumour; cytostatic; lung cancer; human; ds; gene; L548S.

XX Homo sapiens.  
XX  
XX US2003211510-A1.  
XX  
XX 13-NOV-2003.

XX 28-OCT-2002; 2002US-00283017.

XX 30-JUN-1999; 99US-00346492.  
PR 15-OCT-1999; 99US-00419356.  
PR 17-DEC-1999; 99US-00466867.  
PR 30-DEC-1999; 99US-00476300.  
PR 06-MAR-2000; 2000US-00519642.  
PR 22-MAR-2000; 2000US-00533077.  
PR 10-APR-2000; 2000US-00546259.  
PR 27-APR-2000; 2000US-00560406.  
PR 05-JUN-2000; 2000US-00589184.  
PR 11-JUL-2000; 2000US-00614124.  
PR 29-AUG-2000; 2000US-00651563.  
PR 08-SEP-2000; 2000US-00658824.  
PR 26-SEP-2000; 2000US-00671325.  
PR 06-OCT-2000; 2000US-00677419.  
PR 30-OCT-2000; 2000US-00702705.  
PR 13-DEC-2000; 2000US-00736457.  
PR 03-MAY-2001; 2001US-00849626.  
PR 10-JUL-2001; 2001US-00902941.  
PR 29-OCT-2001; 2001US-00017754.  
PR 28-MAR-2002; 2002US-00113872.

XX (CORI-) CORIXA CORP.

XX Henderson RA, Wang T, Watanabe Y, Kalos MD, Sleath PR;  
PI Johnson JC, Retter MW, Durham M, Carter D, Fanger GR, Vedvick TS;  
PI Bangur CS, McNabb A;  
XX WPI; 2004-167010/16.  
DR P-PSDB; ADJ21351.

XX Novel polynucleotide encoding lung tumor polypeptides, useful for  
PT diagnosing, preventing and treating cancer e.g. lung cancer.

XX Example 14; SEQ ID NO 1914; 99pp; English.

XX The invention relates to a novel isolated polynucleotide comprising a  
CC sequence chosen from any one of 40 lung tumour polypeptides or its  
complements, fragments or degenerate variants. The method of the

CC invention has cytostatic applications and may be useful for detecting and  
CC treating lung cancer in a patient, as well as for inhibiting the  
CC development of lung cancer in a patient via incubating CD4+ and/or CD8+ T  
CC cells isolated from a patient with at least one component chosen from a  
CC polypeptide, polynucleotide or antigen presenting cell (APC) of the  
CC invention and administering an effective amount of the proliferated T  
CC cells to the patient. The current sequence is that of the human lung  
CC cancer-related DNA of the invention. The current sequence is not shown in  
CC the specification per se but is available on the USPTO web-site  
CC http://seqdata.uspto.gov/sequence.html?DocID=20030211510.

XX  
SQ Sequence 624 BP; 221 A; 110 C; 178 G; 115 T; 0 U; 0 Other;  
Query Match 2.1%; Score 19; DB 13; Length 624;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 890 CCTTTTTCGGGCAACTTT 908  
|||||  
Db 559 CCTTTTTCGGGCAACTTT 541

Search completed: September 12, 2005, 13:13:17  
Job time : 505.287 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 06:38:26 ; Search time 3957.19 Seconds  
(without alignments)  
11167.303 Million cell updates/sec

Title: US-10-780-002-36  
Perfect score: 912  
Sequence: 1 atcacgctattctcaacaa.....ttttccgggcaactttcatc 912

Scoring table: OLIGO\_NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sce:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID       | Description |
|------------|-------|-------------|--------|----|----------|-------------|
| 1          | 861   | 94.4        | 928    | 8  | AY046015 | Arabidops   |
| 2          | 808   | 88.6        | 808    | 8  | AY142490 | Arabidops   |
| 3          | 777   | 85.2        | 777    | 8  | AY093714 | Arabidops   |
| 4          | 638   | 70.0        | 911    | 8  | AY058115 | Arabidops   |
| 5          | 424   | 46.5        | 57180  | 8  | ATF18D22 | Arabidops   |
| 6          | 140   | 15.4        | 856    | 12 | AY202287 | Arabidops   |
| 7          | 133   | 14.6        | 208    | 11 | BX321957 | Arabidops   |
| 8          | 35    | 3.8         | 160    | 11 | AL773330 | Arabidops   |
| 9          | 23    | 2.5         | 961    | 8  | AF178576 | Rauwolfia   |
| 10         | 23    | 2.5         | 238185 | 2  | AC111417 | Rattus no   |
| 11         | 22    | 2.4         | 143168 | 8  | AP003765 | Oryza sat   |
| 12         | 22    | 2.4         | 146690 | 8  | AP004671 | Oryza sat   |
| 13         | 22    | 2.4         | 163246 | 9  | AC010196 | Homo sapi   |
| 14         | 22    | 2.4         | 187111 | 9  | AC087244 | Homo sapi   |
| 15         | 21    | 2.3         | 31740  | 2  | AC025128 | Homo sapi   |
| 16         | 21    | 2.3         | 39210  | 3  | U41508   | Caenorhabd  |
| 17         | 21    | 2.3         | 101620 | 9  | AC102942 | Homo sapi   |
| 18         | 21    | 2.3         | 139301 | 8  | AP003537 | Oryza sat   |
| 19         | 21    | 2.3         | 143739 | 8  | AC132483 | Oryza sat   |

|    |    |     |        |    |            |                    |
|----|----|-----|--------|----|------------|--------------------|
| 20 | 21 | 2.3 | 147588 | 9  | AC098612   | Homo sapi          |
| 21 | 21 | 2.3 | 161850 | 9  | AF306443   | Homo sapi          |
| 22 | 21 | 2.3 | 235029 | 2  | AC131419   | Rattus no          |
| 23 | 20 | 2.2 | 3819   | 5  | BC070706   | Xenopus l          |
| 24 | 20 | 2.2 | 8592   | 6  | AX346884   | Sequence           |
| 25 | 20 | 2.2 | 12042  | 1  | AB001089   | Arabidops          |
| 26 | 20 | 2.2 | 43351  | 3  | CBRG47121  | Arabidops          |
| 27 | 20 | 2.2 | 54252  | 8  | AB017060   | Arabidops          |
| 28 | 20 | 2.2 | 66805  | 2  | AC101558   | Mus muscu          |
| 29 | 20 | 2.2 | 110000 | 2  | AP006491_3 | Continuation (4 of |
| 30 | 20 | 2.2 | 110121 | 9  | AL590989   | Human DNA          |
| 31 | 20 | 2.2 | 134701 | 5  | BX255885   | Zebrafish          |
| 32 | 20 | 2.2 | 167619 | 5  | AL929045   | Zebrafish          |
| 33 | 20 | 2.2 | 174569 | 9  | AC022533   | Homo sapi          |
| 34 | 20 | 2.2 | 175171 | 10 | AC112158   | Mus muscu          |
| 35 | 20 | 2.2 | 176155 | 2  | AC020879   | Mus muscu          |
| 36 | 20 | 2.2 | 176775 | 2  | AC016978   | Homo sapi          |
| 37 | 20 | 2.2 | 178968 | 9  | AP006274   | Papio ham          |
| 38 | 20 | 2.2 | 192808 | 10 | AL645566   | Mouse DNA          |
| 39 | 20 | 2.2 | 195918 | 10 | AL805935   | Mouse DNA          |
| 40 | 20 | 2.2 | 199885 | 2  | AC116869   | Mus muscu          |
| 41 | 20 | 2.2 | 210472 | 9  | AC035139   | Homo sapi          |
| 42 | 20 | 2.2 | 218807 | 2  | BX511249   | Mus muscu          |
| 43 | 20 | 2.2 | 231764 | 5  | BX005426   | Zebrafish          |
| 44 | 20 | 2.2 | 240691 | 2  | AC108297   | Rattus no          |
| 45 | 20 | 2.2 | 241421 | 2  | AC110427   | Rattus no          |

ALIGNMENTS

|            |   |        |      |        |                 |
|------------|---|--------|------|--------|-----------------|
| RESULT 1   | AY046015  | 928 bp | mRNA | linear | PLN 18-SEP-2002 |
| LOCUS      | AY046015  | 928 bp | mRNA | linear | PLN 18-SEP-2002 |
| DEFINITION | Arabidopsis thaliana putative alpha-hydroxynitrile lyase (AF5910300) mRNA, complete cds.  |        |      |        |                 |
| ACCESSION  | AY046015  |        |      |        |                 |
| VERSION    | AY046015.1  |        |      |        |                 |
| KEYWORDS   | FLI CDNA.   |        |      |        |                 |
| SOURCE     | Arabidopsis thaliana (thale cress)  |        |      |        |                 |
| ORGANISM   | Arabidopsis thaliana  |        |      |        |                 |
| REFERENCE  | 1 (bases 1 to 928)  |        |      |        |                 |
| AUTHORS    | Yanada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K., Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.   |        |      |        |                 |
| TITLE      | Arabidopsis Full Length cDNA Clones   |        |      |        |                 |
| JOURNAL    | Unpublished   |        |      |        |                 |
| REFERENCE  | 2 (bases 1 to 928)  |        |      |        |                 |
| AUTHORS    | Yanada, K., Liu, S.X., Pham, P.K., Banh, J., Dale, J.M., Gibson, H.A., Goldsmith, A.D., Jiang, P.X., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Koesema, E., Lam, B., Lin, J., Meyer, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tracy, S.E., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A. |        |      |        |                 |
| TITLE      | Direct Submission   |        |      |        |                 |
| JOURNAL    | Submitted (06-JUL-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  |        |      |        |                 |
| COMMENT    | RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of AFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.  |        |      |        |                 |

The Salk, Stanford, PGSC (SSP) Consortium members carried out the

sequencing and annotation of the RAFL cDNAs: Yamada, K., Liu, S.X., Pham, P.K., Bath, J., Dale, J.M., Gibson, H.A., Goldsmith, A.D., Jiang, P.X., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koeseema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tracy, S.E., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGECC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGECC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

#### FEATURES

Location/Qualifiers

1..928

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="5"

/clone="RAFL05-09-N03 (R09923)"

/ecotype="Columbia"

/note="this clone is in a modified pBluescript vector (lambda ZAP) as a XhoI/SstI insert."

1..928

/gene="At5g10300"

1..61

/gene="At5g10300"

62..838

/gene="At5g10300"

/codon\_start=1

/evidence="experimental"

/product="putative alpha-hydroxynitrile lyase"

/protein\_id="AAK7689.1"

/db\_xref="GI:15028131"

/translation="MERKHFVLVHNAVGVWYKPLLESAGHRVTAVELAASGI DRPIQAVETVDEYSKPLIETLKSIPENEVILVGFSGGNIALAADI PPAKIKVLV FLNAPLPDTHVPSHLVDKMEMPGGLGDCFSFHSHTNRNLSLKMGPFKMKRLQ NCPIEDYLELAKMLHRQSFPTDLDSKEKFESEYGSQVRVYVMSSEDAIPCFIRW MIDNFNVSKVYIEDGGDHVMVLSKPKQLFDSLAIATDYM"

misc\_difference 590

/gene="At5g10300"

/note="compared to genomic sequence resulting in an amino acid sequence difference"

/replace="g"

839..928

/gene="At5g10300"

3'UTR

ORIGIN

Query Match 94.4%; Score 861; DB 8; Length 928;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 911; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCACGCTATTCTCAACAGCAGAACACCTTAACCAAAAGAAACAAACCTCTACGAGACA 60  
DB 1 ATCACGCTATTCTCAACAGCAGAACACCTTAACCAAAAGAAACAAACCTCTACGAGACA 60

QY 61 TATGAGAGAGAAACATCACTTCGTGTAGTTTCAACACGCTTATCATGAGCGCTGGATCTG 120  
DB 61 TATGAGAGAGAAACATCACTTCGTGTAGTTTCAACACGCTTATCATGAGCGCTGGATCTG 120

QY 121 GTACAGAGCTCAAGCCCTCTTGAATACGCGGCCACCGCTTACTGCTGCGAATCTGCG 180  
DB 121 GTACAGAGCTCAAGCCCTCTTGAATACGCGGCCACCGCTTACTGCTGCGAATCTGCG 180

QY 181 CGCCTCCGGGATCGACCCAGACCAATCCAGCGCTTGAACCGTCGACGATATCTCAA 240  
DB 181 CGCCTCCGGGATCGACCCAGACCAATCCAGCGCTTGAACCGTCGACGATATCTCAA 240

QY 241 ACCGTTGATCGAAACCCCTCAAAATCTCTCCAGAGAACGAGAGGTAATTCGTGTTGGATT 300  
DB 241 ACCGTTGATCGAAACCCCTCAAAATCTCTCCAGAGAACGAGAGGTAATTCGTGTTGGATT 300

QY 301 CAGCTTCGGAGGCATCAACATCGCTCTCGCGCGGACATATTTCCGCGGAGATTAAGGT 360  
DB 301 CAGCTTCGGAGGCATCAACATCGCTCTCGCGCGGACATATTTCCGCGGAGATTAAGGT 360

QY 361 TCTTGTGTCTCTCAACGCTTCTTGGCCCGACACAAACCCAGCTGCTTCTCACGCTTCGGA 420  
DB 361 TCTTGTGTCTCTCAACGCTTCTTGGCCCGACACAAACCCAGCTGCTTCTCACGCTTCGGA 420

QY 421 CAAGTATATGAGATGCTCGAGGTTTGGGAGATGTGAGTTTTCATCTCATGAACAAAG 480  
DB 421 CAAGTATATGAGATGCTCGAGGTTTGGGAGATGTGAGTTTTCATCTCATGAACAAAG 480

QY 481 AAATGGGACGATGAGTTTATTGAGATGGGACCAAAATTCATGAAGGCACCTCTTTACCA 540  
DB 481 AAATGGGACGATGAGTTTATTGAGATGGGACCAAAATTCATGAAGGCACCTCTTTACCA 540

QY 541 AAATTTGTCCTCATAGAGATTACGAGCTGGCAAAATTTGTCATAGGCAAGGTCATTTTT 600  
DB 541 AAATTTGTCCTCATAGAGATTACGAGCTGGCAAAATTTGTCATAGGCAAGGTCATTTTT 600

QY 601 CACAGAGATCTTCAAAAGAAAGAAATTTAGCGAGGAGATATGTTCCGTCGCAACG 660  
DB 601 CACAGAGATCTTCAAAAGAAAGAAATTTAGCGAGGAGATATGTTCCGTCGCAACG 660

QY 661 AGTTTACGTAATGAGTAGTGAAGCAAGCCATCCCTGCGATTTCATTCGTTGGATGAT 720  
DB 661 AGTTTACGTAATGAGTAGTGAAGCAAGCCATCCCTGCGATTTCATTCGTTGGATGAT 720

QY 721 TGATAAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTGATGCT 780  
DB 721 TGATAAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTGATGCT 780

QY 781 CTCCAAAACCCCAAAACCTCTTTGATCTCTCTCTGCTATTCGCCACCGATTAATGTAATA 840  
DB 781 CTCCAAAACCCCAAAACCTCTTTGATCTCTCTCTGCTATTCGCCACCGATTAATGTAATA 840

QY 841 ATCTTAAGTCGGTTTACTTTTTTCTCATCGTTACTTAATAAAACAAACCCCTTTTCCGG 900  
DB 841 ATCTTAAGTCGGTTTACTTTTTTCTCATCGTTACTTAATAAAACAAACCCCTTTTCCGG 900

QY 901 GCAACTTTCATC 912  
DB 901 GCAACTTTCATC 912

RESULT 2  
LOCUS AY142490 808 bp mRNA linear PLN 23-SEP-2002  
DEFINITION Arabidopsis thaliana putative alpha-hydroxynitrile lyase  
ACCESSION AY142490  
VERSION AY142490.1 GI:23296321  
KEYWORDS FLI CDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
REFERENCE 1 (bases 1 to 808)  
AUTHORS Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.  
TITLE Arabidopsis Open Reading Frame (ORF) Clones  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 808)  
AUTHORS Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.  
TITLE Direct Submission

JOURNAL Submitted (16-AUG-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES

source

1. .808

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="5"

/clone="C104719"

/ecotype="Columbia"

/note="This clone is in pUNI 51."

1. .808

/gene="At5g10300"

1. 777

/gene="At5g10300"

/codon\_start=1

/evidence=experimental

/product="putative alpha-hydroxynitrile lyase"

/protein\_id="AA13041.1"

/db\_xref="GI:23296322"

/translations="MERKHFVLVHNAYHGAWIYKLPLESAGHRVTVAVELASGI  
DRPIQAVETVDEYSKPLIETLKSLENBEVILVFGSFGGINIALAADIPPAKIKVLV  
FLNAPLDTHVPSHVDIKMPEGLGCEFSHETRNGLSLAKMGPKFKARLYQ  
NCPIDYELAKMLHROGSPFTEDLSKKERFSEGVGVQVYVMSSEDAIPCFDIRW  
MIDNFNVSKVYEIDGGDHVWMLSKPQLFDSLAIATDM"

778. .808

/gene="At5g10300"

3' UTR

ORIGIN

Query Match 88.6%; Score 808; DB 8; Length 808;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 808; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 ATGGAGGAGAAACATCATTCTGTTAGTTTACAAAGCTTATCATGAGCCTGATCTGG 121

DB 1 ATGGAGGAGAAACATCATTCTGTTAGTTTACAAAGCTTATCATGAGCCTGATCTGG 60

QY 122 TACAAGCTCAAGCCCTCTTGAATCAGCGCGGCACGCGGTACTCTCTCGAATCTGGC 181

DB 61 TACAAGCTCAAGCCCTCTTGAATCAGCGCGGCACGCGGTACTCTCTCGAATCTGGC 120

QY 182 GCCTCCGGGATCGACCGACCAACCAATCCAGCGCGTTGAAACCGTCGACGAATCTCCAA 241

DB 121 GCCTCCGGGATCGACCGACCAACCAATCCAGCGCGTTGAAACCGTCGACGAATCTCCAA 180

QY 242 CGTTGATCGAAACCTCAAATCTCTTCCAGAGAACGAGAGTAATCTGGTTGGATTC 301

DB 181 CGTTGATCGAAACCTCAAATCTCTTCCAGAGAACGAGAGTAATCTGGTTGGATTC 240

QY 302 AGCTTCGGAGGATCAACATCGCTCTCGCGCGGCACATATTTCCGGGAGAGTAAGTT 361

DB 241 AGCTTCGGAGGATCAACATCGCTCTCGCGCGGCACATATTTCCGGGAGAGTAAGTT 300

QY 362 CTGTGTCTCTCAAGCCTCTTCCCGCACACACACCGTCTCTCAAGTTCTCGAC 421

DB 301 CTGTGTCTCTCAAGCCTCTTCCCGCACACACACCGTCTCTCAAGTTCTCGAC 360

QY 422 AAGTATATGAGATGCTCGAGGTTTGGAGATTGTGAGTTTTCATCTCATGAACAAGA 481

DB 361 AAGTATATGAGATGCTCGAGGTTTGGAGATTGTGAGTTTTCATCTCATGAACAAGA 420

QY 482 AATGGACCATGAGTTTATGAGATGGGACCAAAATTCATGAGGACGCTTTACCAA 541

DB 421 AATGGACCATGAGTTTATGAGATGGGACCAAAATTCATGAGGACGCTTTACCAA 480

QY 542 AATTGTCCCATGAGGATTACGAGCTGGCGGCAAAATGTGTATGATGAGGAGGGTCAATTTTC 601

DB 481 AATTGTCCCATGAGGATTACGAGCTGGCGGCAAAATGTGTATGATGAGGAGGGTCAATTTTC 540

QY 602 ACAGAGGATCTATCAAGAGAAAGAAAGTTTAGCGAGGAGGATATGTTTCGTTGCAACGA 661

Db 541 ACAGAGGATCTATCAAGAGAAAGAAAGTTTAGCGAGGAGGATATGTTTCGTTGCAACGA 600

QY 662 GTTTACGTAATAGTAGTAGTGAACAAAGCCATCCCTCGATTTTCATTCGTTGGATGATT 721

Db 601 GTTTACGTAATAGTAGTAGTGAACAAAGCCATCCCTCGATTTTCATTCGTTGGATGATT 660

QY 722 GATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGCGGAGATCACATGTTGATGCTC 781

Db 661 GATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGCGGAGATCACATGTTGATGCTC 720

QY 782 TCCAAACCCCAAAAACCTCTTTCGACTCTCTCTCTGCTATTTGCCACCGATTATATGTAATA 841

Db 721 TCCAAACCCCAAAAACCTCTTTCGACTCTCTCTCTGCTATTTGCCACCGATTATATGTAATA 780

QY 842 TCTTAAGTCGCTTTTACTTTTTTTCAT 869

Db 781 TCTTAAGTCGCTTTTACTTTTTTTCAT 808

RESULT 3

AY093714

LOCUS AY093714

DEFINITION Arabidopsis thaliana AT5g10300/F19D22\_70 mRNA, complete cds.

ACCESSION AY093714

VERSION AY093714.1

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 777)

AUTHORS Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

TITLE Arabidopsis ORF clones

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 777)

AUTHORS Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

TITLE Direct Submission

JOURNAL Submitted (29-MAR-2002) Salk Institute Genomic Analysis Laboratory (SIGnAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGSC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAPL cDNAs: Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L., Chang, E., Dale, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Shinn, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)



|                           |     |   |     |
|---------------------------|-----|---|-----|
|                           |     | /product="AT5g10300/F18D22_70"                              |     |
|                           |     | /protein_id="AAL25532.1"                                    |     |
|                           |     | /db_xref="GI:16648679"                                      |     |
|                           |     | /translation="MERKHFVLVNHAYHGAMIWYKLPILLSAGHRVTAVELASGI     |     |
|                           |     | DRPFOAVDEYISKPLLETLSQENEGILVFGPGINIALAADIFFPAKIKVLV         |     |
|                           |     | FLNAPUDTTHVPSHVLVDKYMMPGGLGDCFFSSHBRTNGTWSLLRMGPKFKALYQ     |     |
|                           |     | NCPIDYELAKMLHRQGSFTEDLSKKEKFSBEGVSGVQVTVMSSEDAIPCDFIRW      |     |
|                           |     | MIDNFNVSKVYEIDGGDHVMLSQPKLPDLSAIATDM"                       |     |
| 3'UTR                     |     | 847. .911   |     |
| ORIGIN                    |     |   |     |
| Query Match               |     | 70.0%; Score 638; DB 8; Length 911;                         |     |
| Best Local Similarity     |     | 99.5%; Pred. No. 0;   |     |
| Matches 838; Conservative |     | 0; Mismatches 4; Indels 0; Gaps 0;                          |     |
| Qy                        | 61  | TATGGAGAGAACATCCTCTCGTGTAGTTTCAACAGCTTATCATGGAGCCTGGATCTG   | 120 |
| Db                        | 69  | TATGGAGAGAAACATCCTCTCGTGTAGTTTCAACAGCTTATCATGGAGCCTGGATCTG  | 128 |
| Qy                        | 121 | GTACAAGCTCAAGCCCTCTCTGAATCAGCCGCCACCGCTTACTGCTGTGAACTCGC    | 180 |
| Db                        | 129 | GTACAAGCTCAAGCCCTCTCTGAATCAGCCGCCACCGCTTACTGCTGTGAACTCGC    | 188 |
| Qy                        | 181 | CGCTCCGGATCGAACCCACGCAATCAGGCCGTTGAAACCGTGCAGCAATACTCCAA    | 240 |
| Db                        | 189 | CGCTCCGGATCGAACCCACGCAATCAGGCCGTTGAAACCGTGCAGCAATACTCCAA    | 248 |
| Qy                        | 241 | ACGCTTGATCGAAACCTCNAATCTCTCCAGAGAACGAGGTAATTCGTGGATT        | 300 |
| Db                        | 249 | ACGCTTGATCGAAACCTCNAATCTCTCCAGAGAACGAGGTAATTCGTGGATT        | 308 |
| Qy                        | 301 | CAGCTTCGAGGATCAACATCGCTCTCGCCGCCGACATATTTCCGGCGAAGATTAAAGT  | 360 |
| Db                        | 309 | CAGCTTCGAGGATCAACATCGCTCTCGCCGCCGACATATTTCCGGCGAAGATTAAAGT  | 368 |
| Qy                        | 361 | TCCTGTGTTCTCAACGCCCTTCTGCGCGACACAACCCACGTCCTTCTCAGCTTCTGA   | 420 |
| Db                        | 369 | TCCTGTGTTCTCAACGCCCTTCTGCGCGACACAACCCACGTCCTTCTCAGCTTCTGA   | 428 |
| Qy                        | 421 | CAAGTATATGAGATCGCTGGAGTTTGGAGATTGTGAGTTTCACTCATGAACACAG     | 480 |
| Db                        | 429 | CAAGTATATGAGATCGCTGGAGTTTGGAGATTGTGAGTTTCACTCATGAACACAG     | 488 |
| Qy                        | 481 | AAATGGACGATGAGTTTATTGAAGATGGACCAAAATTCATGAAGCACGCTTTACCA    | 540 |
| Db                        | 489 | AAATGGACGATGAGTTTATTGAAGATGGACCAAAATTCATGAAGCACGCTTTACCA    | 548 |
| Qy                        | 541 | AAATTGCTCCATAGAGATTACGAGTCGGCAAAAATGTTGCATAGCAAGGTCATTTT    | 600 |
| Db                        | 549 | AAATTGCTCCATAGAGATTACGAGTCGGCAAAAATGTTGCATAGCAAGGTCATTTT    | 608 |
| Qy                        | 601 | CACAGAGATCTATCAAGAGAAAGTTTACGAGGAGAGATATGTTCCGTCGCAACG      | 660 |
| Db                        | 609 | CACAGAGATCTATCAAGAGAAAGTTTACGAGGAGAGATATGTTCCGTCGCAACG      | 668 |
| Qy                        | 661 | AGTTTACGTAATGAGTAGTGAAGACAAGCCATCCCTCGCATTTTCATTCGTTGATGAT  | 720 |
| Db                        | 669 | AGTTTACGTAATGAGTAGTGAAGACAAGCCATCCCTCGCATTTTCATTCGTTGATGAT  | 728 |
| Qy                        | 721 | TGATAATTTCAAGCTCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTATGCT | 780 |
| Db                        | 729 | TGATAATTTCAAGCTCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTATGCT | 788 |
| Qy                        | 781 | CTCCAAACCCCAAAATCTTTGATCTCTCTGCTATGTCACCGCATATATGTAATA      | 840 |
| Db                        | 789 | CTCCAAACCCCAAAATCTTTGATCTCTCTGCTATGTCACCGCATATATGTAATA      | 848 |
| Qy                        | 841 | ATCTTAAGTCGTTTTTACTTTTTCTCATCGTTACTATAAACAACCCCTTTTCCGG     | 900 |
| Db                        | 849 | ATCTTAAGTCGTTTTTACTTTTTCTCATCGTTACTATAAACAACCCCTTTTCCGG     | 908 |
| Qy                        | 901 | GC 902  |     |

|  |  |           |              |
|--|--|-----------|--------------|
| Db   | 909                                      | GC 910    |              |
| RESULT 5   |  |           |              |
| ATF18D22   |  |           |              |
| LOCUS  |  |           |              |
| DEFINITION   |  |           |              |
| ATF18D22   | 57180 bp                                 | DNA       | linear       |
| Arabidopsis thaliana   | DNA chromosome 5,                        | BAC clone | F18D22 (BSSA |
| project).  |  |           |              |
| ACCESSION  |  |           |              |
| AL360334   |  |           |              |
| VERSION  |  |           |              |
| AL360334.1   | GI:8953404                               |           |              |
| KEYWORDS   |  |           |              |
| SOURCE   |  |           |              |
| ORGANISM   |  |           |              |
| Arabidopsis thaliana   | (thale cress)                            |           |              |
| Arabidopsis thaliana   |  |           |              |
| Eukaryota; Viridiplantae;  | Streptophyta; Embryophyta; Tracheophyta; |           |              |
| Spermatophyta; Magnoliophyta;  | eudicotyledons; core eudicots;           |           |              |
| rosids; eurosids II; Brassicales;  | Brassicaceae; Arabidopsis.               |           |              |
| 1 (bases 1 to 52804)   |  |           |              |
| Bevan, M., Zimmermann, W., Gruenewald, A., Wambutt, R., Kalicki, J.,     |  |           |              |
| Wohlmann, P., Smith, A., Bancroft, I., Mewes, H.W., Rudd, S., Lemcke, K. |  |           |              |
| and Mayer, K.F.X.  |  |           |              |
| Unpublished  |  |           |              |
| 2 (bases 52233 to 57180)   |  |           |              |
| Bevan, M., Voelckert, G., Grymoprez, B., Voet, M., Robben, J.,           |  |           |              |
| Bancroft, I., Mewes, H.W., Rudd, S., Lemcke, K. and Mayer, K.F.X.        |  |           |              |
| Unpublished  |  |           |              |
| 3 (bases 1 to 57180)   |  |           |              |
| EU Arabidopsis sequencing, project.                                      |  |           |              |
| Direct Submission  |  |           |              |
| Submitted (05-JUL-2000)  | MIPS, at the Max-Planck-Institut fuer    |           |              |
| Biochemie, Am Klopferplatz 18a, D-82152 Martinsried, FRG, E-mail:        |  |           |              |
| lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de                    |  |           |              |
| Project  |  |           |              |
| Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge        |  |           |              |
| Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,         |  |           |              |
| E-mail: michael.bevan@bbsrc.ac.uk  |  |           |              |
| Information on performance of analysis and a more detailed               |  |           |              |
| annotation of this entry and other sequences of chromosomes 3, 4         |  |           |              |
| and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.       |  |           |              |
| FEATURES   |  |           |              |
| Location/Qualifiers  |  |           |              |
| 1..57180   |  |           |              |
| /organism="Arabidopsis thaliana"   |  |           |              |
| /mol_type="Genomic DNA"  |  |           |              |
| /variety="Columbia"  |  |           |              |
| /db_xref="taxon:3702"  |  |           |              |
| /chromosome="5"  |  |           |              |
| misc_feature   |  |           |              |
| 1..20186   |  |           |              |
| /note="Overlap to BAC T31P16, please refer to                            |  |           |              |
| EMBL:AL356332 for analysis and annotation"                               |  |           |              |
| 17729..21213   |  |           |              |
| /gene="F18D22_10"  |  |           |              |
| complement(join(17729..17941,18038..18145,18237..18323,                  |  |           |              |
| 18409..18489,18699..18833,18917..19138,19223..19303,                     |  |           |              |
| 19389..19550,19645..19739,19924..20019,20122..20263,                     |  |           |              |
| 20453..20548,20735..20873,21134..21213))                                 |  |           |              |
| /gene="F18D22_10"  |  |           |              |
| complement(join(17729..17941,18038..18145,18237..18323,                  |  |           |              |
| 18409..18489,18699..18833,18917..19138,19223..19303,                     |  |           |              |
| 19389..19550,19645..19739,19924..20019,20122..20263,                     |  |           |              |
| 20453..20548,20735..20873,21134..21213))                                 |  |           |              |
| /gene="F18D22_10"  |  |           |              |
| /note="Contains Glutamine amidotransferases class-II                     |  |           |              |
| active site AAL-6"   |  |           |              |
| /codon_start=1   |  |           |              |
| /product="asparagine synthetase ASN3"                                    |  |           |              |
| /protein_id="CAB96680.1"   |  |           |              |
| /db_xref="GI:8953405"  |  |           |              |
| /db_xref="GOA:Q9LFU1"  |  |           |              |
| /db_xref="UniProt/TrEMBL:Q9LFU1"   |  |           |              |
| /translation="WCGILAVLCVDNSQAKRSRIIEIIRLRHGRPDWSGLHCVEDC                 |  |           |              |
| YLAHERLAIVDPTSGDQDPLYNEDKTIATVNGEYINHKALRENKSHQFRTGSDCEVI                |  |           |              |
| AHLYEEGEEFVMDLGMFAVLIDTRDKSFIAARDAIGITPLYIGWGLDGSVWFASE                  |  |           |              |
| MKALSDCECEPCFPFGHIYSKQGLRWYNNPFWSEVPSTPYDPLVVRNTPEKAV                    |  |           |              |
| IKRLMTDVPFGVLLSGGLDSSLVASLRHLEKSEACQWGSKLHTFCIGIKGSPDLK                  |  |           |              |
| AGREVADYLGTRHRLHFTVQGDIDAIEVIYHYVDVTTIRASTPFLMSRKIKSL                    |  |           |              |

GKVMLSGGSDEIFGGYLYPHKAPNKKFEHETCRKIKALHOYDCLRANKSTSAWGV  
EARVPFLDKEFTINVAMSIDPEWKMIKRPDLGRIEKWVLRNAPFDDKKNPFLPKHILYRQK  
EQFSQGVGSWIDGUKDRANKRVSETMLMNASFVFPDNTPLTKERAYITRTIFPEKFFPK  
SAARATVPGGSPVACSTAKAVEMDAWSONLDPSGRAALGVHRVSAYGEDKTEDSRPEK  
LQKLAETKPAIV"  
complement (17729. .17941)  
/gene="F18D22\_10"  
/number=1  
complement (17942. .18037)  
/gene="F18D22\_10"  
/number=1  
complement (18038. .18145)  
/gene="F18D22\_10"  
/number=2  
complement (18146. .18236)  
/gene="F18D22\_10"  
/number=2  
complement (18237. .18323)  
/gene="F18D22\_10"  
/number=3  
complement (18324. .18408)  
/gene="F18D22\_10"  
/number=3  
complement (18409. .18489)  
/gene="F18D22\_10"  
/number=4  
complement (18490. .18698)  
/gene="F18D22\_10"  
/number=4  
complement (18699. .18833)  
/gene="F18D22\_10"  
/number=5  
complement (18834. .18916)  
/gene="F18D22\_10"  
/number=5  
complement (18917. .19138)  
/gene="F18D22\_10"  
/number=6  
complement (19139. .19222)  
/gene="F18D22\_10"  
/number=6  
complement (19223. .19303)  
/gene="F18D22\_10"  
/number=7  
complement (19304. .19388)  
/gene="F18D22\_10"  
/number=7  
complement (19389. .19550)  
/gene="F18D22\_10"  
/number=8  
complement (19551. .19644)  
/gene="F18D22\_10"  
/number=8  
complement (19645. .19739)  
/gene="F18D22\_10"  
/number=9  
complement (19740. .19923)  
/gene="F18D22\_10"  
/number=9  
complement (19924. .20019)  
/number=10  
complement (20020. .20121)  
/gene="F18D22\_10"  
/number=10  
complement (20122. .20263)  
/gene="F18D22\_10"  
/number=11  
complement (20264. .20452)  
/gene="F18D22\_10"  
/number=11  
complement (20453. .20548)  
/gene="F18D22\_10"

intron  
complement (20549. .20734)  
/gene="F18D22\_10"  
/number=12  
complement (20735. .20873)  
/gene="F18D22\_10"  
/number=13  
complement (20874. .21133)  
/gene="F18D22\_10"  
/number=13  
complement (21134. .21213)  
/gene="F18D22\_10"  
/number=14  
21823. .24163  
/gene="F18D22\_20"  
complement (join(21823. .22209, 22312. .22440, 22508. .22712,  
22838. .23739, 23877. .23946, 24033. .24163))  
/gene="F18D22\_20"  
complement (join(21823. .22209, 22312. .22440, 22508. .22712,  
22838. .23739, 23877. .23946, 24033. .24163))  
/gene="F18D22\_20"  
/note="strong similarity to non-phototropic hypocotyl 3 -  
Arabidopsis thaliana, EMBL:AF180390  
Contains Prokaryotic membrane lipoprotein lipid attachment  
site AA304-314"  
/codon\_start=1  
/product="non-phototropic hypocotyl 3-like protein"  
/protein\_id="CAB96681.1"  
/db\_xref="GI:8953406"  
/db\_xref="GOA:Q9LFU0"  
/db\_xref="UniProt/TsEMBL:Q9LFU0"  
/translation="MNSVSAARLKSFQDAGPVCNKSIIPSRVVTILANSPEKKDRSW  
YVKSIQPTDLSIQVNDITFKAKHPLISKCYISSIELKSTSENGYHLKLEFPGGA  
WRTFTILKFCYNLPDLNPLNVAIPRCASELYIMTEFEAGNLLISFEATFTFVVLAS  
DIDTLVLRSCNLSIPWAEVLQIVRCDLLAKWACNDNIPEDVDVDRNERCLYNDIA  
TLNDRMEESNLGCOBKETIESLVLPQSGAVSCHFLRLMLKTSIVVSASPALI  
SVNKRVMALDANVCDLLIPNFKNEEQOERVIPEFFLMHEQOQVLKPSISKLLD  
NYLAEIAKDPYLPITKQVLAEMLPENAWKCHDGLYRAIDMFLKTHSPDHRRRLC  
KTMNCEKLSLDACLHAAQNDRPLRTIVQINTQVLFSEQVKNRMMDQLKPEKEENS  
GGREDKMSRDNEIITKLKELENVKKMSLQSDYNELOQOYERLSSKQKSHNWGL  
RWQVKKSFTQKREDETRERTRRRSTGORTSPRRMSMS"  
complement (21823. .22209)  
/gene="F18D22\_20"  
/number=1  
complement (22210. .22311)  
/gene="F18D22\_20"  
/number=1  
complement (22312. .22440)  
/gene="F18D22\_20"  
/number=2  
complement (22441. .22507)  
/gene="F18D22\_20"  
/number=2  
complement (22508. .22712)  
/gene="F18D22\_20"  
/number=3  
complement (22713. .22837)  
/gene="F18D22\_20"

Query Match 46.5%; Score 424; DB 8; Length 57180;  
Best Local Similarity 100.0%; Pred. No. 1.9e-227; Mismatches 0; Indels 0; Gaps 0;  
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATCAGCGTATTCTCAACAAGCAAGAACCTTAAACCAAAAAAGAAACCTCTACGAGACA 60  
Db 44418 ATCAGCGTATTCTCAACAAGCAAGAACCTTAAACCAAAAAAGAAACCTCTACGAGACA 44477  
QY 61 TATGGAGAGGAACATCACTTCGTGTAGTTTCACACGCTTATCATGAGCGCTGGATCTG 120  
Db 44478 TATGGAGAGGAACATCACTTCGTGTAGTTTCACACGCTTATCATGAGCGCTGGATCTG 44537  
QY 121 GTACAAGCTCAAGCCCTCTCTTGAATCAGCGCGGCACCGCGTTACTGTCTGCAACTCGC 180



```

Db 44538 GTACAAAGCTCAAGCCCTCCTTGAATCAGCGGCCACCGGTTACTGCTGTCGAATCTGC 44597
|
QY 181 CGCCTCGGGATCGACCCAGCACCAATCCAGGCCGTTGAAACCGTCGACGAATACTCCAA 240
|
Db 44598 CGCCTCGGGATCGACCCAGCACCAATCCAGGCCGTTGAAACCGTCGACGAATACTCCAA 44657
|
QY 241 ACCGTTGATCGAAACCCCTCAAATCTCTTCCAGAGAACGAAAGAGTAATTCCTGGTTGGATT 300
|
Db 44658 ACCGTTGATCGAAACCCCTCAAATCTCTTCCAGAGAACGAAAGAGTAATTCCTGGTTGGATT 44717
|
QY 301 CAGCTTCGGAGGATCAACATCGCTCTCGCCCGCGACATATTTCCGGCGAAGATTAAGGT 360
|
Db 44718 CAGCTTCGGAGGATCAACATCGCTCTCGCCCGCGACATATTTCCGGCGAAGATTAAGGT 44777
|
QY 361 TCTGTGTCTCTCAACGCTTCTTGGCCGACACCAACCCAGTCGCTTCTCAGCTTCTGGA 420
|
Db 44778 TCTGTGTCTCTCAACGCTTCTTGGCCGACACCAACCCAGTCGCTTCTCAGCTTCTGGA 44837
|
QY 421 CAAG 424
|
Db 44838 CAAG 44841
|
RESULT 6
AY202287 856 bp DNA linear SYN 27-MAY-2003
LOCUS Arabidopsis thaliana sequence flanking Ds3 end of Ds-Trap insertion
DEFINITION from line GT6478.
ACCESSION AY202287
VERSION AY202287.1 GI:27898241
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM
REFERENCE 1 (bases 1 to 856)
AUTHORS May B.P., Simorowski J., Arroyo J.-M., Vaughn, M.W., Shen, R.,
McCombie, W.R. and Martienssen, R.A.
TITLE Arabidopsis genomic sequences flanking Ds enhancer and gene traps
in transgenic lines
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 856)
AUTHORS May B.P., Simorowski J., Arroyo J.-M., Vaughn, M.W., Shen, R.,
McCombie, W.R. and Martienssen, R.A.
TITLE Direct Submission
JOURNAL Submitted (17-DEC-2002) Cold Spring Harbor Laboratory, 1 Bungtown
Road, Cold Spring Harbor, NY 11724, USA
COMMENT http://genetraps.cshl.org.
FEATURES
source
Location/Qualifiers
1..856
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/clone="GT6478.Ds3.05.08.98.b.888"
/ecotype="Landsberg erecta"
/transgenic
/notes="transgenic line GT6478"
1..34
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
repeat_region <1..34
/transposon="transposon Ds"
ORIGIN
Query Match 15.4%; Score 140; DB 12; Length 856;
Best Local Similarity 99.0%; Pred. No. 5.7e-67;
Matches 290; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 555 AGGATTACGAGCTGGCAAAATGTTGCATAGGCAAGGTCATTTTTCACAGAGGATCTAT 614
|

```

```

Db 336 AGGATTACGAGCTGGCAAAATGTTGCATAGGCAAGGTCATTTTTCACAGAGGATCTAT 395
|
QY 615 CAAAGAAAGAAAGTTTACGAGGAGGATATGGTTCCGTCGCAACGAGTTTACGTAATGA 674
|
Db 396 CAAAGAAAGAAAGTTTACGAGGAGGATATGGTTCCGTCGCAACGAGTTTACGTAATGA 455
|
QY 675 GTAGTGAAGACAAAGCCATCCCTCGCATTTTCATTCTGTTGGATGATTAATTTTCAACG 734
|
Db 456 GTAGTGAAGACAAAGCCATCCCTCGCATTTTCATTCTGTTGGATGATTAATTTTCAACG 515
|
QY 735 TCTCGAAGTCTACGAGATCGATGGCGGAGATCACATGCTGCTCTCAACCCCAAA 794
|
Db 516 TCTCGAAGTCTACGAGATCGATGGCGGAGATCACATGCTGCTCTCAACCCCAAA 575
|
QY 795 AACTCTTTGACCTCTCTCTGCTATTGCCACCGATTATATGTAATAATCTTAA 847
|
Db 576 AACTCTTTGACCTCTCTCTGCTATTGCCACCGATTATATGTAATAATCTTAA 628
|
RESULT 7
BX321957 208 bp DNA linear STS 10-JUN-2003
LOCUS Arabidopsis thaliana transposon insertion STS SM_3.35356, sequence
DEFINITION tagged site.
ACCESSION BX321957
VERSION BX321957.1 GI:29420086
KEYWORDS STS; STS, sequence tagged site.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM
REFERENCE 1
AUTHORS Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCullagh, B.,
Murphy, G., Langham, S., LeGrys, C., Jones, J.D.G. and Bevan, M.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 208)
AUTHORS Clarke, J.H.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-2003) Clarke J.H., John Innes Centre, Colney
Lane, Norwich, NR4 7UJ, UK
COMMENT AT denotes an activation tag dissociation transposon within a
single line, Et an enhancer trap dissociation transposon, GT a gene
trap dissociation transposon, MT a mis-expression enhancer trap
dissociation transposon, SM a defective suppressor mutator
transposon. 3 denotes a sequence derived from the 3' end of the
transposon, 5 denotes a sequence derived from the 5' end of the
transposon BBSRC GARNET, ARIS project
On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock
code: N122067.
FEATURES
Location/Qualifiers
1..208
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia-0 NASC stock code N1092"
/db_xref="taxon:3702"
/clone="AL360334"
/notes="Derived from superpool 26.07 NASC code Unknown"
1..208
/standard_name="SM_3.35356"
STS
1..208
ORIGIN
Query Match 14.6%; Score 133; DB 11; Length 208;
Best Local Similarity 100.0%; Pred. No. 4.8e-63;
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 780 TCTCCAAACCCCAAAACTCTTTGACTCTCTCTGCTATTGCCACCGATTATGTAAT 839
|
Db 1 TCTCCAAACCCCAAAACTCTTTGACTCTCTCTGCTATTGCCACCGATTATGTAAT 60
|
QY 840 AATCTTAAGTCGGTTTACTTTTCTCATCGTTACTAATAAAACAAACCCCTTTTCG 899
|
Db 61 AATCTTAAGTCGGTTTACTTTTCTCATCGTTACTAATAAAACAAACCCCTTTTCG 120
|

```

```

Qy 900 GGCAACTTTCATC 912
Db 121 GGCAACTTTCATC 133

RESULT 8
AL773330/c
LOCUS
DEFINITION Arabidopsis thaliana transposon insertion STS SM_3.1688, sequence
tagged site.
ACCESSION AL773330
VERSION AL773330.1 GI:21535517
KEYWORDS STS; STS, sequence tagged site.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1
Clarke,J.H., Bowles,B., Carter,J., Hart,D., McCullagh,B.,
Murphy,G., Langham,S., LeGrys,C., Jones,J.D.G. and Bevan,M.
Unpublished
2 (bases 1 to 160)
Clarke,J.H.
Direct Submission
Submitted (18-JUN-2002) Clarke J.H., John Innes Centre, Colney
Lane, Norwich, NR4 7UJ, UK
AT denotes an activation tag dissociation transposon within a
single line, ET an enhancer trap dissociation transposon, GT a gene
trap dissociation transposon, MT a mis-expression enhancer trap
dissociation transposon, SM a defective suppressor mutator
transposon. 3 denotes a sequence derived from the 3'end of the
transposon, 5 denotes a sequence derived from the 5'end of the
transposon BESRC GARNET, ATIS project
On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock
code: N56946.
FEATURES
source
Location/Qualifiers
1..160
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia-0 NASC stock code N1092"
/db_xref="taxon:3702"
/clone="AL360334"
/note="Derived from superpool 5.20 NASC code N40219"
1..160
/standard_name="SM_3.1688"

STS
1..160
/standard_name="SM_3.1688"

ORIGIN
Query Match 3.8%; Score 35; DB 11; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 555 AGGATTACGAGTGGCAAAAATGTCATAGCAA 589
Db 35 AGGATTACGAGTGGCAAAAATGTCATAGCAA 1

RESULT 9
AF178576
LOCUS
DEFINITION Rauwolfia serpentina polynuridine aldehyde esterase mRNA, complete
cds.
ACCESSION AF178576
VERSION AF178576.1 GI:6651392
KEYWORDS Rauwolfia serpentina (serpentwood)
ORGANISM Rauwolfia serpentina
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Gentianales; Apocynaceae; Rauvolfioideae;
Vinceae; Rauvolfia.
1 (bases 1 to 961)

```

```

AUTHORS Dogru,E., Warzecha,H., Seibel,F., Haebel,S., Lottepeich,F. and
Stockigt,J.
TITLE The gene encoding polynuridine aldehyde esterase of monoterpeneoid
indole alkaloid biosynthesis in plants is an ortholog of the
alpha/betahydrolase super family
Eur. J. Biochem. 267 (5), 1397-1406 (2000)
20156243
JOURNAL MEDLINE
PUBMED 10691977
REFERENCE 2 (bases 1 to 961)
AUTHORS Warzecha,H., Dogru,E. and Stoeckigt,J.
TITLE Direct Submission
JOURNAL Submitted (18-AUG-1999) Institut fur Pharmazie, Johannes Gutenberg
Universitaet, Staudinger Weg 5, Mainz 55099, Germany
FEATURES
source
Location/Qualifiers
1..961
/organism="Rauwolfia serpentina"
/mol_type="mRNA"
/db_xref="taxon:4060"
/note="Cell suspension culture"
15..809
/function="catalyzes 16-epi-vellosimine formation from
polynuridine aldehyde"
/function="involved in ajmaline biosynthesis"
/codon_start=1
/product="polynuridine aldehyde esterase"
/protein_id="AAF22288.1"
/db_xref="GI:6651393"
/translation="MHSAANAKQKQHFVLVHGCGLGAWIYKLPLESAHGKVTAVD
LSAAGINPRRLDEIHTPRDYSEPLMEYMASIPDPKVVLLGHSPGMSGLAMILTYPE
KISAVFMSAMPDPNHSLTVPFEKYNEKCPADMLDSQFSTYGNPENGMSMLIGPO
PMALKMFQNCSEVLELAKMLTRPGSLFFODLAKAKKESRYGSKRAYIFCNEKDS
PFVFQKWFVESGADKVKKEADHGMLSQPREVCKCLLDISDS"

ORIGIN
Query Match 2.5%; Score 23; DB 8; Length 961;
Best Local Similarity 100.0%; Pred. No. 0.79;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 113 TGGATCTGTTACAGCTCAAGCC 135
Db 84 TGGATCTGTTACAGCTCAAGCC 106

RESULT 10
AC111417/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-33309, *** SEQUENCING IN PROGRESS
ACCESSION AC111417
VERSION AC111417.4 GI:24942305
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 238185)
Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Ayodeji,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswato,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Chen,A.,
Chacko,J., Chavez,D., Chen,G., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederick,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,

```



Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoques, M., Hollins, B., Howells, S., Hulvik, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhera, L., Louised, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mathewney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoelameh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Prannkoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Reigh, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission  
Unpublished  
2 (bases 1 to 238185)  
Worley, K. C.

Direct Submission  
Submitted (19-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 238185)  
Rat Genome Sequencing Consortium.

Direct Submission  
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Nov 13, 2002 this sequence version replaced gi:23603976.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GMGO  
Center clone name: CH230-33309  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 226513 bases at least Q40  
Consensus quality: 228567 bases at least Q30  
Consensus quality: 229730 bases at least Q20

Estimated insert size: 234005; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 4 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 231322: contig of 231322 bp in length  
\* 231323 231422: gap of unknown length  
\* 231423 235627: contig of 4205 bp in length  
\* 235628 235727: gap of unknown length  
\* 235728 236737: contig of 1010 bp in length  
\* 236738 236837: gap of unknown length  
\* 236838 238185: contig of 1348 bp in length.

FEATURES  
source  
1..238185  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-33309"  
59308..60745  
/note="wgs\_contig"  
231423..232441  
/note="wgs\_contig"

misc\_feature  
misc\_feature  
231423..232441  
/note="wgs\_contig"

ORIGIN  
Query Match 2.5%; Score 23; DB 2; Length 238185;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 452 GATTGTGAGTTTTCATCTCATGA 474  
|||||  
DB 135161 GATTGTGAGTTTTCATCTCATGA 135139  
|||||

RESULT 11  
AP003765  
LOCUS  
DEFINITION  
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 7,  
BAC clone:OJ1003\_C06.  
ACCESSION  
AP003765.5 GI:50509923  
VERSION  
AP003765  
KEYWORDS  
SOURCE  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
REFERENCE  
1  
Sasaki, T., Matsumoto, T. and Yamamoto, K.  
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC  
clone:OJ1003\_C06  
Published Only in Database (2001)  
2 (bases 1 to 143168)  
Sasaki, T., Matsumoto, T. and Yamamoto, K.  
Direct Submission  
Submitted (20-JUN-2001) Takuji Sasaki, National Institute of  
Agrobiological Sciences, Rice Genome Research Program; Kamondai  
2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
(E-mail: [tsasaki@nias.affrc.go.jp](mailto:tsasaki@nias.affrc.go.jp), URL: <http://rgp.dna.affrc.go.jp/>,  
Tel: 81-298-38-7441, Fax: 81-298-38-7468)  
On Jul 22, 2004 this sequence version replaced gi:34393713.  
Genes were predicted from the integrated results of the following:  
GENSCAN (<http://CCR-081.mit.edu/GENSCAN.html>), FGENESH  
(<http://www.softberry.com/>), GeneMark.hmm  
(<http://opal.biology.gatech.edu/GeneMark/>), GlimmerM  
([http://www.tigr.org/tdb/glimmerm/glmr\\_form.html](http://www.tigr.org/tdb/glimmerm/glmr_form.html)), RiceHMM

(http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor (http://bioinformatics.iastate.edu/cgi-bin/ep.cgi), sim4 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The genomic sequence was searched against NCBI NonRedundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP or DDBJ. Protein homologies of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no.

A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence.

The orientation of the sequence is from -21M13 to M13rev of the BAC clone. This sequence of OJ1003\_C06 clone has an overlap with P0524G08 clone (DDBJ: AP004671) at 5' end and an overlap with OJ1458\_B07 (DDBJ: AP003832) at 3' end. The sequence was generated by combining Monanto and RGP-Japan sequencing data. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rgp.dna.affrc.go.jp/Genomeseq.html>.

FEATURES  
source

1. .143168  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="7"  
/clone="OJ1003\_C06"  
complement (1942..4991)  
/gene="OJ1003\_C06.101"  
complement (join(<1942..2007,2095..2154,2618..2658,3821..3854,3935..3985,4926..>4991))  
/gene="OJ1003\_C06.101"  
/note="start and end point are not identified"  
complement (join(1942..2007,2095..2154,2618..2658,3821..3854,3935..3985,4926..4991))  
/gene="OJ1003\_C06.101"  
/note="contains EST(s): AU030351 (E50932), AU030350 (E50932)"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="BAD30245.1"  
/db\_xref="GI:50509924"  
/translation="WDERFRRLMLFLPVVRSDYCAESGSSKSGTQSSRAQEAETRGNKSSSDQFMKLLKMWAEKKVGATKAELEFCKTFEAKHKLVTKEINLDAORFLNAYE KGS"  
complement (6337..6991)  
/gene="OJ1003\_C06.102"  
complement (join(<6337..6615,6764..>6991))  
/gene="OJ1003\_C06.102"  
/note="start and end point are not identified"  
complement (join(6337..6615,6764..6991))  
/gene="OJ1003\_C06.102"  
/note="predicted by FGENESH etc."  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAD30246.1"  
/db\_xref="GI:50509925"  
/translation="WADRRPAYNTTAKRLQRCRPHGGHDKDAIAHLRKMWFPPKKE NSPRKRYRPOEGYDTRCSRCPGDHNPAASSLHPTAPPLRAGPGHLRPATR RPPOVRPKENPDWTVSLRPPLELLTSSRLNPTETPHKEGEPFKGEGADWQOQGRRTAASSHEH"  
complement (8772..9955)

/gene="OJ1003\_C06.103"  
complement (join(<8772..8959,9611..9796,9898..>9955))  
/gene="OJ1003\_C06.103"  
/note="start and end point are not identified"  
complement (join(8772..8959,9611..9796,9898..9955))  
/gene="OJ1003\_C06.103"  
/note="predicted by GeneMark.hmm etc."  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAD30247.1"  
/db\_xref="GI:50509926"  
/translation="MEQTGSSGRQOGLAKGLHDNAVYVMKGVDSDEKAVLAVDM ANKRLQIVSVDAERIYDDRDYSYTSQTSYFTTAAGVKGKLRPLKFHAQYPHKRL GETLSRDNPMDLHEPQLQIDGFLGKTDETFSDNPMDLE"  
complement (10031..10462)  
/gene="OJ1003\_C06.104"  
complement (<10031..>10462)  
/gene="OJ1003\_C06.104"  
/note="start and end point are not identified"  
complement (10031..10462)  
/gene="OJ1003\_C06.104"  
/note="predicted by GeneMark.hmm etc."  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAD30248.1"  
/db\_xref="GI:50509927"  
/translation="MATVFPYRSRTEHLTLRPPCHGGDPLDAHQPLPWLDDVRAYI ADNSTTAADLGNHLYRPTAAGILHLCLVHTKTKVITLDEGQGEVAFIDLKRGI NPHQARPMLSWSKATAARHQLSNELQVLIALPGCHR"  
complement (12401..20516)  
/gene="OJ1003\_C06.105"  
complement (join(<12401..12654,12751..12825,12909..12988,13056..13136,13227..13483,13578..13806,13889..14043,14125..14244,14351..14429,14514..14650,14780..14882,14973..15246,15337..15487,18354..18509,18621..18711,19872..19911,20211..>20516))  
/gene="OJ1003\_C06.105"  
/note="supported by full-length cDNA(s): AK100505"  
complement (12401..20516)  
/gene="OJ1003\_C06.105"  
/note="contains full-length cDNA(s): AK100505  
non-coding transcript  
probably inactive due to including stop codon(s) in CDS"  
complement (20915..23256)  
/gene="OJ1003\_C06.106"  
complement (join(20915..21640,22598..22819,22922..23256))  
/gene="OJ1003\_C06.106"  
/note="supported by full-length cDNA(s): AK058442"  
complement (join(20985..21640,22598..22819,22922..23219))  
/gene="OJ1003\_C06.106"  
/note="contains full-length cDNA(s): AK058442"  
/codon\_start=1  
/product="lipase-like protein"  
/protein\_id="BAD30249.1"  
/db\_xref="GI:50509928"  
/translation="MATRPARLSAPVAAILVAALCAATASAAVNTTTATAPSGG CYSHLFTGNSLIDTGNFIHYSTSPGPVARSFYGETFRRTGRNSGRLIVDFIVER LGFPYWPYLAGRSERDFRYGAFVAGSTALNQLLKKGLSVAGITPYSLAVQVGY PKYVLAMLASTEQERKEAMARSVEFMGDFGNDYHLPLFQNKTLWVRPLVPRVYRI AGAVEELVGLGATTVPGLPFLCGVPRLLFLFRDGGAGDRDPATGCLRGNDGLAAL HNALLRRLAELRAHPGVTTAYADYGEWELVSNPAASGDFDALTACCAAGGPPYNG NNTVHCSDPGATQCAQDFSRISDGLHMTAVYRIMARGVLDDGFPADPPIMSRCHGY 24414..26326  
/gene="OJ1003\_C06.107"  
join(<24414..24473,24569..24666,25426..25525,25698..25880,26204..>26326)  
/gene="OJ1003\_C06.107"  
/note="start and end point are not identified"  
join(24414..24473,24569..24666,25426..25525,25698..25880,26204..26326)  
/gene="OJ1003\_C06.107"  
/note="contains EST(s): C74320 (E30875), AU088598 (C53863), AU166063 (E30875), C97298

```

(C53863)"
/codon_start=1
/product="transport protein particle component Bet3-like
protein"
/protein_id="BAD30250.1"
/db_xref="GI:50509929"
/translation="MPPLTPGKSGDALFGSVERVNAELFTLYGAIVRQLLFDLBEVE
EVNKQIDQMGYNIGTRLVDFELAKNSVRCVDFKFTADVIAGLFGFMFLGVATVTNW
DAEGTSCSVLENDPLVDFVELPDTCCQGLVCYNLGVIRGALEVMVKTEVTWVRDM
LRGDDAYENRVKLTQVPEEYPKDD"
27936..31529
/gene="OJ1003_C06.108"
/genes="OJ1003_C06.108"
/notes="start and end point are not identified"
join(27936..29276,29457..29769,30718..31529)
/gene="OJ1003_C06.108"
/codon_start=1
/product="putative RUSH-1alpha"
/protein_id="BAD30251.1"
/db_xref="GI:50509930"
/translation="WASTSRGSGEDDDPEYLLGFIIVANVLOYVGGRAITRRS
VGLVREPHNPHDANIRVDNARGEKIGHGRRRAALAPLLDAGHVAAHGLVPRPAS
KRLYSLPCQVHLFARPPHAALAAASGIDLIIHVDHPEFALSSEAIHQEQTKRSR
2.4%; Score 22; DB 8; Length 143168;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 CTCGCCCTCCGGGATGACC 197
Db 28317 CTCGCCCTCCGGGATGACC 28338

RESULT 12
AP004671
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 7,
AP004671
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1
Sasaki.T., Matsumoto.T. and Yamamoto.K.
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, PAC
Clone:P0524G08
Published Only in Database (2002)
2 (bases 1 to 146690)
Sasaki.T., Matsumoto.T. and Yamamoto.K.
Direct Submission
Submitted (23-JAN-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kamondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Jul 29, 2002 this sequence version replaced gi:18307749.
Genes were predicted from the integrated results of the following:
GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
(http://www.softberry.com/), GeneMark.hmm
(http://opal.biology.gatech.edu/GeneMark/), GlimmerM
(http://opal.biology.gatech.edu/glimmer/glmr form.html), RiceHMM
(http://rgp.dna.affrc.go.jp/RiceHMM/), _SplicePredictor
(http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4
(http://globin.cse.psu.edu/html/docs/sim4.html), gap2
(http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The
genomic sequence was searched against NCBI NonRedundant Protein
database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
sequence database at RGP or DDBJ. Protein homologies of the coding

```

```

regions were searched against NCBI NonRedundant Protein database
with BLASTP. ESTs represent the identified cDNA sequences using
BLASTN with the corresponding DDBJ accession no. and RGP clone ID.
Full-length cDNAs represent the identified cDNA sequences using
BLASTN with the corresponding DDBJ accession no.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with full-length cDNA or
EST homology (covering almost the entire length of partial
sequence) is classified as an 'unknown' protein. A gene predicted
by two or more gene prediction programs is classified as a
'hypothetical' protein according to IRGSP standard. A gene
predicted by a single gene prediction program is also classified as
a probable 'hypothetical' protein and is included as a
miscellaneous feature of the sequence.
The orientation of the sequence is from SP6 to T7 of the PAC clone.
This sequence of P0524G08 clone has an overlap with OJ1340_C08
(DDBJ: AP005292) clone at 5' end and with OJ1003_C06 (DDBJ:
AP003765) at 3' end. Detailed information on overlap and assembly
quality together with annotation of this entry is available at
http://rgp.dna.affrc.go.jp/GenomeSeq.html.
Location/Qualifiers
1..146690
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="7"
/clone="P0524G08"
1573..3882
/gene="P0524G08.101"
join(1573..1933,3403..3841)
/notes="supported by full-length cDNA (s): AK059275"
join(1590..1933,3403..3882)
/gene="P0524G08.101"
/notes="supported by full-length cDNA (s): AK070897"
join(1721..1933,3403..3585)
/gene="P0524G08.101"
/notes="contains Set's (s): AU162548 (R3914), AU032339 (R3914)
contains full-length cDNA (s): AK059275, AK070897
similar to Arabidopsis thaliana chromosome 3, At3g12260"
/codon_start=1
/product="unknown protein"
/protein_id="BAC16465.1"
/db_xref="GI:23237891"
/translation="MAFTMRAVKVPPNSASLEEARHVRVDFFRQACRAIPSIIMEIYNL
DDVTPSQLRSTIAKEIRKNOGVNPKVIDMLFKGMEELGTEHAKQRHVIQYV
VGQKGLVQDMKQDQSSDFLKFFYTSNYS"
4758..9355
/gene="P0524G08.102"
join(4758..5101,6083..6734,6851..6939,7398..7462,
7571..7626,7711..7831,7920..7994,8293..8346,8434..8528,
8608..8782,8897..9355)
/gene="P0524G08.102"
/notes="supported by full-length cDNA (s): AK070161"
join(5040..5101,6083..6734,6851..6939,7398..7462,
7571..7626,7711..7831,7920..7994,8293..8346,8434..8528,
8608..8782,8897..9141)
/gene="P0524G08.102"
/notes="contains full-length cDNA (s): AK070161"
/codon_start=1
/product="putative xylose kinase"
/protein_id="BAC16466.1"
/db_xref="GI:23237892"
/translation="WVGRRSLPEGSFLGDSSTQSLKATVNLNELLIVASDIVNFD
DLPHYKTEGGVRRPADDDHIFPSITMVEAELELLEKLPKINFSKVAVLSGSGQOH
GSVTTWKSGHVALSLSDPTKLLSQLKDAFTMDSPIMWDSSTTKQCREISVUGGAL
ELSKLTGSRAVERFTGRIKYRTVFPQVDDTERISLVSVFASILVGNVAFIDNCL
EGANMLADINQRTWSKTVEATPGLBEKGLAPAVAGRIAPFYFVRQFIDNCL
VIOWSGDNPNSLAGLTNLTGDLAISLTGTSVFTAEAKPSLEGHVFPNPVPDGY
MVMLCYKNGSLTDRVNSCAEKSWDFNSYLEKTPPLNGKGLGFPYKDHILPLPV

```

GFHRVIVENLNDVTSNNLVERVEFDPPESEVRATIEIQQLLSMRCHAEFQWPNPKR  
LIATGASNERILHSIAQIFCCPVFTVQRPDSASLGAALRAAGLWLCNBSGSPVPI  
CMYQNLKTSLSGAKLAVSTGEGVEDKELLEKYTVLMRKRMEIERRLVEKIGRA"  
complement(10645..14070)  
/gene="P0524G08.103"  
complement(join(<10645..10870,11911..12051,12128..12186,  
12255..12561,13086..13167,13244..13377,13462..13546,  
13626..13738,13823..13902,14011..>14070))  
/gene="P0524G08.103"  
/note="start and end point are not identified"  
complement(join(10645..10870,11911..12051,12128..12186,  
12255..12561,13086..13167,13244..13377,13462..13546,  
13626..13738,13823..13902,14011..14070))  
/gene="P0524G08.103"  
/note="predicted by GeneMark.hmm etc.  
similar to Oryza sativa chromosome 8, OJ1081\_B12.12"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAC16467.1"  
/db\_xref="GI:23237893"  
/translation="WASASTPTIFSIDGVILDEBGTMMQFTIPKKLENEFRPSLTGLC  
VITKVLHLEGRSGSIPNYGFEFCFPDQIPSKSRISKELIGVISHVGYDYAGKTSK  
KNRKLIRKDEQEGIVLWGEYGSFDEAFVLOKSTDHKIWAIVLGLTAGTYLTKT  
EATSLATQILYFDSITRIVEVQSRVYKPLATIKDELGTLDVAFVNVVAEDLVEVNAVTO  
ATQNKIDATEHAIALDTAIGKTRLFHAMTKYSSHTFINVYKLSKSPVENNTSLM  
LPTLENTKAKESATQKATDEGLTMEHYSKEDQHPITPPSLQPPETTLNDNTVNI  
IPSAKRALQFKEHLIDQPRNHEVQLIHNTYLAQVSKIASSTAISITLQKGSSTD  
LSKLQTPQAHVVMYMKLPHMHIIIDGENVKNEI"  
complement(14288..18600)  
/gene="P0524G08.104"  
complement(14288..18600)  
/gene="P0524G08.104"  
/note="probably inactive due to including frameshift(s) in  
CDS  
pseudogene,helicase"  
/pseudo  
25168..27249  
/gene="P0524G08.105"  
join(<25168..25348,26073..26110,26865..27006,  
27095..>27249)  
/gene="P0524G08.105"  
/note="start and end point are not identified"  
join(25168..25348,26073..26110,26865..27006,27095..27249)  
/gene="P0524G08.105"  
/note="predicted by GENSCAN etc."  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAC16468.1"  
/db\_xref="GI:23237894"  
/translation="WSRNDKSRAAVDRIKTAALSAAGLNRSQAKHAATSAAARNVPM  
GRRKQDAAGRGRLILAAIILVLEIHEAPLOFAGRAQANFAVHQLFVATARSV  
GGIDGMGMSAAHMTASRSGFRAGRLIILADGSGQDEBTSELTSESTSSAWILC  
PWLLPSWIDD"  
30279..30467  
/gene="P0524G08.106"  
30279..30467  
/gene="P0524G08.106"  
/note="start and end point are not identified"  
30279..30467  
/gene="P0524G08.106"  
/note="predicted by GlimmerM etc."  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAC16469.1"  
/db\_xref="GI:23237895"  
/translation="WEGPRAEAAAGRSIDGWLWTASLIVLALLAALMYQQPYAPWPV  
QPCNCKPACVRQSRMTW"  
35625..36011  
/gene="P0524G08.107"  
35625..36011  
/gene="P0524G08.107"  
/note="supported by full-length cDNA(s) : AK107877"  
35625..36011

/gene="P0524G08.107"  
/note="contains full-length cDNA(s) : AK107877  
non-coding transcript  
probably inactive due to including stop codon(s) in CDS"  
join(37684..37833,39835..40872)  
/gene="P0524G08.108-1"  
join(37684..37833,39835..40872)  
/gene="P0524G08.108-2"  
/note="supported by full-length cDNA(s) : AK102159"  
join(37687..38060,39726..40877)  
/gene="P0524G08.108-1"  
join(37687..38060,39726..40877)  
/gene="P0524G08.108-1"  
/note="supported by full-length cDNA(s) : AK063952"  
join(37782..38060,39726..40379)  
/gene="P0524G08.108-2"  
/note="contains full-length cDNA(s) : AK063952,AK099354"  
Query Match 2.4%; Score 22; DB 8; Length 146690;  
Best Local Similarity 100.0%; Pred. No. 3.6; Mismatches 0; Gaps 0;  
Matches 22; Conservative 0;  
Qy 176 CTCGCGCCTCCGGATCGACC 197  
|||||  
Db 99064 CTCGCGCCTCCGGATCGACC 99085  
RESULT 13  
AC010196 163246 bp DNA linear PRI 28-SEP-2002  
LOCUS Homo sapiens 12 BAC RP11-202H2 (Roswell Park Cancer Institute Human  
BAC Library) complete sequence.  
AC010196  
AC010196.5 GI:6139072  
HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (Bases 1 to 163246)  
AUTHORS Muzny,D.M., Adams,C., Bailey,M., Barbara,J., Blankenburg,K.,  
Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,  
Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,  
David,R., Delgado,O., Deshazo,D., Ding,Y., Donah-Rashid,N.,  
Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,  
Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,  
Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,  
Holloway,C., Hosak,H., Issar,A., Jackson,L.E., Jackson,L., Jia,Y.,  
Jones,M., Kelly,S., Kneitz,S., Kondejewski,N., Kong,Y., Kovar,C.,  
Lau,S., Leal,B., Lee,E., Li,Z., Lichtarge,O., Liu,J., Liu,W.,  
Logan,O., Lu,J., Lucier,R., Marondel,I., Martin,R., Martinez,C.,  
McLeod,M.P., Mei,G., Merscher,S., Miller,A., Montgomery,K.T.,  
Morgan,M., Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N.,  
Nguyen,S., Osval,G., Parish,B., Paxton,S., Payton,B., Perez,L.,  
Pu,L.L., Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J.,  
Scherer,S., Shah,E., Shen,H., Shim,C., Simon,M., Sparks,A.,  
Stamps,A., Sucgang,R., Tabor,P., Taylor,T., Vasquez,L., Vinson,R.,  
Vo,Q., Wahbah,M., Watlington,S., Weinstock,G., Weinstock,I.R.,  
Williamson,A., Worley,K., Wren,J., Wrensford,G., Xiang,A.M.,  
Yang,R., Yu,W., Zhou,X., Kucherlapati,R., Nelson,D. and Gibbs,R.  
Direct Submission  
Unpublished  
JOURNAL 2 (Bases 1 to 163246)  
REFERENCE Worley,K.C.  
AUTHORS Direct Submission  
TITLE Submitted (15-SEP-1999) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
JOURNAL 3 (Bases 1 to 163246)  
REFERENCE Worley,K.C.  
AUTHORS Direct Submission  
TITLE Submitted (29-OCT-1999) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Baylor Plaza, Houston, TX 77030, USA  
Worley, K.C.  
Direct Submission  
Submitted (31-MAY-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
5 (bases 1 to 163246)  
Worley, K.C.  
Direct Submission  
Submitted (28-SEP-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Oct 29, 1999 this sequence version replaced gi:6087859.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
[gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Submitted (28-SEP-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Oct 29, 1999 this sequence version replaced gi:6087859.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
[gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

COMMENT

CLONE LENGTH: This sequence does not necessarily represent the  
entire insert of this clone. Overlapping regions of clones are only  
sequenced and submitted once, so the sequence for the remainder of  
the insert may be found in the record for the adjacent clones.  
Overlapping clones are noted at the beginning and end of the  
Features listing.

ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches  
of a local database that includes entries from dbSTS, GDB, and  
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,  
unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST  
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the  
EST and cDNA sequences. Genes demonstrate at least two exons  
flanked by consensus splice sites that maintained sequence  
continuity across the splice junctions. Sequences that are not  
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum  
standard of double strand coverage with a minimum of 2 clones and 2  
reads with no ambiguities or 2 chemistries with a minimum of 2  
clones and 3 reads with no ambiguities. If the sequence quality for  
a region does not meet this standard, it will be indicated in the  
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality  
standards - estimated error rate less than 1 per 10,000 bases.  
Reports of lowest quality individual bases and measures of base  
quality are listed below. Description of the metrics can be found  
at URL:  
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

----- Summary Statistics -----  
Contig length: 80737  
Phrap values in estimate: 80641  
Average error rate (BCM-Phrap estimate): 8.31934e-05  
Fraction of Phrap values less than 40 : 0.02247  
Number of consensus changing edits: 0  
Number of N's in consensus : 3

----- Consensus changing edits -----  
Position Original+Context Edited+Context

----- Distribution of Quality < 40 Bases -----

|      |   |
|------|---|
| 1000 | * |
| 900  | * |
| 800  | * |
| 700  | * |

bases 600 \*  
500 \*  
400 \*  
300 \*  
200 \*  
100 \*  
0 \* \* \* \* \*  
-----  
5 10 15 20 25 30 35 40  
Phrap Value Range

Version: 1.01 qxfo.  
Location/Qualifiers  
1..163246  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/clone="RP11-202H2"  
complement(1..2001)  
/note="overlaps bases 1..2001 of clone AC128681"  
/function="clone overlap"  
5715..5859  
repeat\_region /rpt\_family="MLT1B1"  
5860..6155  
repeat\_region /rpt\_family="AluJb"  
6156..6222  
repeat\_region /rpt\_family="MLT1B1"  
6223..6738  
repeat\_region /rpt\_family="MLT1D"  
6739..6817  
repeat\_region /rpt\_family="MLT1B1"  
complement(7107..7419)  
repeat\_region /rpt\_family="L1MB1"  
complement(7451..7662)  
repeat\_region /rpt\_family="L1MEC"  
7814..7841  
repeat\_region /rpt\_family="C(AAATA)n"  
complement(7842..8283)  
repeat\_region /rpt\_family="L1MA6"  
8281..9425  
repeat\_region /rpt\_family="L1MA6"  
9493..9514  
repeat\_region /rpt\_family="AT\_rich"  
9577..9809  
repeat\_region /rpt\_family="MIR"  
9812..9834  
repeat\_region /rpt\_family="AT\_rich"  
10805..10845  
repeat\_region /rpt\_family="AT\_rich"  
complement(10846..11182)  
repeat\_region /rpt\_family="L1PA7"  
11183..11231  
repeat\_region /rpt\_family="AT\_rich"  
11753..12148  
repeat\_region /rpt\_family="MLT1A2"  
12166..12244  
repeat\_region /rpt\_family="AT\_rich"  
12265..12554  
repeat\_region /rpt\_family="AluJb"  
12555..12669  
repeat\_region /rpt\_family="GA\_rich"  
complement(12805..13216)  
repeat\_region /rpt\_family="MLT2E"  
complement(13709..14622)  
repeat\_region /rpt\_family="L1MB1"  
complement(14623..15117)  
repeat\_region /rpt\_family="MER9"  
complement(15118..15490)  
repeat\_region /rpt\_family="L1MB1"  
15534..15867  
repeat\_region

```
/rpt_family="L1M4"
17689..17724
/rpt_family="(TTTG)n"
20420..20441
/rpt_family="AT rich"
complement(21000..21182)
/rpt_family="MIR"
complement(21916..22025)
/rpt_family="L2"
23794..24307

Query Match          2.4%; Score 22; DB 9; Length 163246;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 278 GAAGAGGTAATCTCGTTGGAT 299
|||||
Db 113646 GAAGAGGTAATCTCGTTGGAT 113667

RESULT 14
AC087244/c
LOCUS
DEFINITION Homo sapiens 12 BAC RP11-444N1 (Roswell Park Cancer Institute Human
BAC Library) complete sequence.
ACCESSION AC087244
VERSION AC087244.17 GI:13470125
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Binage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,
Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,X., Dinh,H.H., Douchwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D.,
Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J.,
Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W.,
Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K.,
Hart,M., Haviak,P., Hawes,A., Hernandez,J., Hernandez,O.,
Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Homsif.,
Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karleson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Louiseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R.,
Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P.,
Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G.,
Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M.,
Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokweto,S.,
Ogulu,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C.,
Shoshitari,N., Sleson,I., Sodergren,E., Sonaiker,T., Sparks,A.,
Stanley,H., Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A.,
Tamerisa,K., Tang,H., Tanson,J., Taylor,C., Taylor,T., Telford,B.,
Thomas,N., Thomas,S., Umani,K., Vasquez,L., Vera,V., Villalon,D.,
Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R.,
Washington,C., Watlington,S., Williams,G., Williamson,A.,
Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J.,
```

Zorrilla,S., Kucherlapati,R. and Gibbs,R.  
Direct Submission  
Unpublished  
2 (bases 1 to 187111)  
Worley,K.C.  
Direct Submission  
Submitted (20-DEC-2000) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 187111)  
Worley,K.C.  
Direct Submission  
Submitted (28-MAR-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
4 (bases 1 to 187111)  
Worley,K.C.  
Direct Submission  
Submitted (15-MAR-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
5 (bases 1 to 187111)  
Worley,K.C.  
Direct Submission  
Submitted (02-MAY-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
6 (bases 1 to 187111)  
Worley,K.C.  
Direct Submission  
Submitted (02-AUG-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
7 (bases 1 to 187111)  
Worley,K.C.  
Direct Submission  
Submitted (03-SEP-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Mar 28, 2001 this sequence version replaced gi:13443049.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
[gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the  
entire insert of this clone. Overlapping regions of clones are only  
sequenced and submitted once, so the sequence for the remainder of  
the insert may be found in the record for the adjacent clones.  
Overlapping clones are noted at the beginning and end of the  
Features listing.

ANNOTATION OF FEATURES:  
STGs are identified using ePCR (Genome Res. 7:541-550) searches  
of a local database that includes entries from dbSTS, GDB, and  
local mapping efforts.  
Repeats are identified using RepeatMasker (A. Smit and P. Green,  
unpublished.) for Human and Mouse sequences.  
Genes and Region of sequence similarity are identified by BLAST  
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the  
EST and cDNA sequences. Genes demonstrate at least two exons  
flanked by consensus splice sites that maintained sequence  
continuity across the splice junctions. Sequences that are not  
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum  
standard of double strand coverage with a minimum of 2 clones and 2  
reads with no ambiguities or 2 chemistries with a minimum of 2  
clones and 3 reads with no ambiguities. If the sequence quality for  
a region does not meet this standard, it will be indicated in the  
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality  
standards - estimated error rate less than 1 per 10,000 bases.  
Reports of lowest quality individual bases and measures of base

quality are listed below. Description of the metrics can be found at URL:  
http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

## QUALSTAT-REPORT.

```

FEATURES
source      Location/Qualifiers
1. .187111
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="12"
   /clone="RP11-444N1"
   complement(3. .238)
   /rpt_family="L1MB8"
   complement(230. .371)
   /rpt_family="L1MB8"
   complement(2548. .2861)
   /rpt_family="L1MB4"
   3075. .3176
   /rpt_family="L2"
   3654. .3753
   /rpt_family="L2"
   4503. .4734
   /rpt_family="MIR"
   4950. .4997
   /rpt_family="(CAA)n"
   5275. .5340
   /rpt_family="MIR"
   complement(6876. .7173)
   /rpt_family="AluJo"
   8248. .8280
   /rpt_family="L1PA13"
   8281. .8678
   /rpt_family="L1PA13"
   8693. .8922
   /rpt_family="L1MB6"
   8923. .9222
   /rpt_family="AluSg"
   9223. .9592
   /rpt_family="L1MB6"
   9782. .9853
   /rpt_family="L2"
   9978. .10106
   /rpt_family="MER115"
   complement(10107. .10750)
   /rpt_family="L1MB3"
   10751. .10882
   /rpt_family="MER115"
   11732. .11798
   /rpt_family="L1MA4"
   11805. .11825
   /rpt_family="AT_rich"
   12087. .12109
   /rpt_family="AT_rich"
   12452. .12628
   /rpt_family="L1MA9"
   12629. .12670
   /rpt_family="(CA)n"
   12671. .12894
   /rpt_family="L1MA9"
   12895. .12922
   /rpt_family="AT_rich"
   complement(13539. .13588)
   /rpt_family="L2"
   complement(15205. .15428)
   /rpt_family="MIR"
   15883. .15903
   /rpt_family="(T)n"
   complement(15904. .16035)
   /rpt_family="FLAM_C"
   18104. .18127
   /rpt_family="AT_rich"
   18531. .18634

```

```

repeat_region  /rpt_family="AT_rich"
                complement(19682. .19747)
                /rpt_family="L2"

```

```

Query Match      2.4%; Score 22; DB 9; Length 187111;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 68 AGGAAACATCACTTCGTGTAG 89
    |||||
Db 80631 AGGAAACATCACTTCGTGTAG 80610

```

## RESULT 15

```

AC025128/c      31740 bp      DNA      linear      HTG 13-JUL-2000
LOCUS      Homo sapiens clone RP11-307E16, LOW-PASS SEQUENCE SAMPLING.
DEFINITION      AC025128
ACCESSION      AC025128
VERSION      AC025128.1 GI:7158939
KEYWORDS      HTG; HTGS PHASE0.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 31740)
AUTHORS      Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE      Homo sapiens, clone RP11-307E16
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 31740)
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
              Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
              Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
              Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
              Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
              Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
              Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
              Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
              Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
              Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
              Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
              McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R.,
              Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
              Murphy,T., Naylor,T., Norman,C.H., O'Connor,T., O'Donnell,P.,
              O'Neil,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,
              Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
              Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
              Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
              Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
              Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
              Young,G., Zainoun,J., Zimmer,A. and Zody,M.
DIRECT SUBMISSION

```

## TITLE

```

JOURNAL      Submitted (05-MAR-2000) Whitehead Institute/MIT Center for Genome
              Research, 320 Charles Street, Cambridge, MA 02141, USA
              All repeats were identified using RepeatMasker:
              Smit, A.F.A. & Green, P. (1996-1997)
              http://ftp.genome.washington.edu/RM/RepeatMasker.html

```

## COMMENT

```

----- Genom Center

```

```

Center: Whitehead Institute/ MIT Center for Genome Research

```

```

Center code: W1BR

```

```

Web site: http://www-seq.wi.mit.edu

```

```

Contact: sequence_submissions@genome.wi.mit.edu

```

```

----- Project Information

```

```

Center project name: L7794

```

```

Center Clone name: 307_E_16

```

```

-----

```

```

* NOTE: This record contains 36 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone

```



\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1 799: contig of 799 bp in length  
\* 800 899: gap of 100 bp  
\* 1693: contig of 794 bp in length  
\* 1694 1793: gap of 100 bp  
\* 1794 2579: contig of 786 bp in length  
\* 2580 2679: gap of 100 bp  
\* 2680 3436: contig of 757 bp in length  
\* 3437 3536: gap of 100 bp  
\* 3537 4294: contig of 758 bp in length  
\* 4295 4394: gap of 100 bp  
\* 4395 5145: contig of 751 bp in length  
\* 5146 5245: gap of 100 bp  
\* 5246 6045: contig of 800 bp in length  
\* 6046 6145: gap of 100 bp  
\* 6146 6962: contig of 817 bp in length  
\* 6963 7062: gap of 100 bp  
\* 7063 7843: contig of 781 bp in length  
\* 7844 7943: gap of 100 bp  
\* 7944 8728: contig of 785 bp in length  
\* 8729 9582: gap of 100 bp  
\* 9583 9582: contig of 754 bp in length  
\* 9583 9582: gap of 100 bp  
\* 9583 10452: contig of 770 bp in length  
\* 10453 10552: gap of 100 bp  
\* 10553 11353: contig of 801 bp in length  
\* 11354 11453: gap of 100 bp  
\* 11454 12246: contig of 793 bp in length  
\* 12247 12347: gap of 100 bp  
\* 12347 13149: contig of 802 bp in length  
\* 13149 13248: gap of 100 bp  
\* 13249 14028: contig of 780 bp in length  
\* 14029 14128: gap of 100 bp  
\* 14129 14900: contig of 772 bp in length  
\* 14901 15000: gap of 100 bp  
\* 15001 15794: contig of 794 bp in length  
\* 15795 15894: gap of 100 bp  
\* 15895 16701: contig of 807 bp in length  
\* 16702 16801: gap of 100 bp  
\* 16802 17587: contig of 786 bp in length  
\* 17588 17687: gap of 100 bp  
\* 17688 18473: contig of 786 bp in length  
\* 18474 18573: gap of 100 bp  
\* 18574 19354: contig of 781 bp in length  
\* 19355 19454: gap of 100 bp  
\* 19455 20237: contig of 783 bp in length  
\* 20238 20337: gap of 100 bp  
\* 20338 21119: contig of 782 bp in length  
\* 21120 21219: gap of 100 bp  
\* 21220 21995: contig of 776 bp in length  
\* 21996 22095: gap of 100 bp  
\* 22096 22872: contig of 777 bp in length  
\* 22873 22972: gap of 100 bp  
\* 22973 23749: contig of 777 bp in length  
\* 23750 23849: gap of 100 bp  
\* 23850 24643: contig of 794 bp in length  
\* 24644 24743: gap of 100 bp  
\* 24744 25561: contig of 818 bp in length  
\* 25562 25661: gap of 100 bp  
\* 25662 26424: contig of 763 bp in length  
\* 26425 26524: gap of 100 bp  
\* 26525 27317: contig of 793 bp in length  
\* 27318 27417: gap of 100 bp  
\* 27418 28208: contig of 791 bp in length  
\* 28209 28308: gap of 100 bp  
\* 28309 29105: contig of 797 bp in length  
\* 29106 29205: gap of 100 bp  
\* 29206 29995: contig of 790 bp in length  
\* 29996 30095: gap of 100 bp  
\* 30096 30893: contig of 798 bp in length  
\* 30894 30993: gap of 100 bp

\* 30994 31740: contig of 747 bp in length.  
FEATURES  
source  
1. .31740  
location/Qualifiers

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="RP11-307E16"  
/clone\_lib="RPC1-11 Human Male BAC"

ORIGIN

Query Match 2.3%; Score 21; DB 2; Length 31740;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 584 AGGCAAGGGTCATTTTCACA 604  
|||  
Db 15333 AGGCAAGGGTCATTTTCACA 15313

Search completed: September 12, 2005, 15:37:32  
Job time : 3966.19 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 12, 2005, 20:27:23 ; Search time 183 Seconds  
(without alignments)

2324.765 Million cell updates/sec

Title: US-10-780-002-2

Perfect score: 1368

Sequence: 1 MKBGFHVLVHGACHGWSW.....MLCEPQKLCASLLTAHKYN 260

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DSV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10780002/runat\_11092005\_104748\_26625/app\_query.fasta\_1.455  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10780002@cgn 1 1 105 @runat\_11092005\_104748\_26625 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfilee1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description         |
|------------|-------|-------------|--------|----|---------------------|
| 1          | 585   | 42.8        | 1078   | 3  | US-08-981-256A-4    |
| 2          | 141   | 10.3        | 840    | 4  | US-09-328-352-1832  |
| 3          | 127   | 9.3         | 795    | 3  | US-09-134-001C-2755 |
| 4          | 126   | 9.2         | 807    | 4  | US-10-107-096-1     |
| 5          | 123   | 9.0         | 4331   | 4  | US-09-710-279-3711  |
| 6          | 120.5 | 8.8         | 2101   | 1  | US-08-106-761-1     |
| 7          | 117   | 8.6         | 1107   | 4  | US-09-902-540-4129  |
| 8          | 117   | 8.6         | 9564   | 4  | US-09-902-540-1026  |
| 9          | 114   | 8.3         | 8654   | 1  | US-08-920-812-6     |
| 10         | 114   | 8.3         | 8654   | 1  | US-08-920-827-6     |
| 11         | 114   | 8.3         | 8654   | 1  | US-08-921-177-6     |
| 12         | 114   | 8.3         | 8654   | 1  | US-08-362-577C-6    |

|    |       |     |         |   |                      |                    |
|----|-------|-----|---------|---|----------------------|--------------------|
| 13 | 114   | 8.3 | 8654    | 2 | US-08-920-828-6      | Sequence 6, Appli  |
| 14 | 113   | 8.3 | 657     | 4 | US-09-252-991A-9351  | Sequence 9351, Ap  |
| 15 | 113   | 8.3 | 855     | 4 | US-09-252-991A-9422  | Sequence 9422, Ap  |
| 16 | 113   | 8.3 | 1101    | 4 | US-09-252-991A-9455  | Sequence 9455, Ap  |
| 17 | 109   | 8.0 | 855     | 4 | US-09-328-352-1853   | Sequence 1853, Ap  |
| 18 | 108   | 7.9 | 1116    | 4 | US-09-602-777A-225   | Sequence 225, App  |
| 19 | 105   | 7.7 | 891     | 4 | US-09-902-540-6739   | Sequence 6739, Ap  |
| 20 | 105   | 7.7 | 4009    | 4 | US-09-902-540-541    | Sequence 541, App  |
| 21 | 103   | 7.5 | 522     | 4 | US-09-637-746-9      | Sequence 9, Appli  |
| 22 | 103   | 7.5 | 1167    | 1 | US-08-232-519-1      | Sequence 1, Appli  |
| 23 | 103   | 7.5 | 1167    | 1 | US-08-456-956-1      | Sequence 1, Appli  |
| 24 | 102   | 7.5 | 834     | 4 | US-09-252-991A-4372  | Sequence 4372, Ap  |
| 25 | 100.5 | 7.3 | 873     | 4 | US-09-248-796A-2815  | Sequence 2815, Ap  |
| 26 | 99    | 7.2 | 1897    | 4 | US-09-637-746-1      | Sequence 1, Appli  |
| 27 | 99    | 7.2 | 2181    | 4 | US-09-248-796A-3259  | Sequence 3259, Ap  |
| 28 | 96.5  | 7.1 | 2772    | 4 | US-09-540-236-1917   | Sequence 1917, Ap  |
| 29 | 96.5  | 7.1 | 269223  | 4 | US-09-596-002-41     | Sequence 41, Appli |
| 30 | 96    | 7.0 | 49272   | 1 | US-08-614-770A-1     | Sequence 1, Appli  |
| 31 | 95.5  | 7.0 | 4403765 | 3 | US-09-103-840A-2     | Sequence 2, Appli  |
| 32 | 95.5  | 7.0 | 4411529 | 3 | US-09-103-840A-1     | Sequence 1, Appli  |
| 33 | 94    | 6.9 | 885     | 3 | US-09-424-349A-1     | Sequence 1, Appli  |
| 34 | 93    | 6.8 | 522     | 4 | US-09-637-746-7      | Sequence 7, Appli  |
| 35 | 93    | 6.8 | 735     | 4 | US-09-902-540-6017   | Sequence 6017, Ap  |
| 36 | 93    | 6.8 | 1272    | 4 | US-09-252-991A-14082 | Sequence 14082, A  |
| 37 | 93    | 6.8 | 1329    | 4 | US-09-252-991A-14036 | Sequence 14036, A  |
| 38 | 93    | 6.8 | 1620    | 4 | US-09-252-991A-14186 | Sequence 14186, A  |
| 39 | 93    | 6.8 | 11029   | 4 | US-09-902-540-1004   | Sequence 1004, Ap  |
| 40 | 92    | 6.7 | 360     | 4 | US-08-956-171E-1291  | Sequence 1291, Ap  |
| 41 | 92    | 6.7 | 360     | 4 | US-08-781-986A-1291  | Sequence 1291, Ap  |
| 42 | 92    | 6.7 | 1796    | 1 | US-07-816-283-11     | Sequence 11, Appli |
| 43 | 92    | 6.7 | 1796    | 1 | US-08-417-103-11     | Sequence 11, Appli |
| 44 | 91    | 6.7 | 1316    | 4 | US-09-461-325-91     | Sequence 91, Appli |
| 45 | 91    | 6.7 | 1316    | 4 | US-10-012-542-91     | Sequence 91, Appli |

ALIGNMENTS

RESULT 1  
US-08-981-256A-4  
; Sequence 4, Application US/08981256A  
; Patent No. 6046042  
; GENERAL INFORMATION:  
; APPLICANT: Mainhard HASSLACHER et al.  
; TITLE OF INVENTION: (S)-HYDROXY-NITRILE-LYASE FROM HEVEA BRASILIENSIS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.  
; STREET: 2033 K Street, N.W., Suite 800  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/981,256A  
; FILING DATE: December 22, 1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew Jacob  
; REGISTRATION NUMBER: 25,154  
; REFERENCE/DOCKET NUMBER: 1553-OZ1112  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 721-8200  
; TELEFAX: (202) 721-8250  
; INFORMATION FOR SEQ ID NO: 4:



```

Db      733 TTGCAGTTGGAA-----CAGCGCTTCGTAGAGATGCGAGGACATTTA 777
               |||::: |||::: |||::: |||:::
Qy      240 AlaMetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIle 255
               :::: |||::: |||::: |||:::
Db      778 AGTTATGTCGACCAACCGCAGCATTTAACCAAAATTAATGTTGACGATT 825
               :::: |||::: |||::: |||:::

RESULT 3
US-09-134-001C-2755
; Sequence 2755, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2755
; LENGTH: 795
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2755

Alignment Scores:
Pred. No.: 1.48e-06 Length: 795
Score: 127.00 Matches: 62
Percent Similarity: 40.47% Conservative: 42
Best Local Similarity: 24.12% Mismatches: 97
Query Match: 9.28% Indels: 56
DB: 3 Gaps: 12

US-10-780-002-2 (1-260) x US-09-134-001C-2755 (1-795)
Qy      3 GluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTyrIlys 22
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      73 GAAGGCATCCCATCATTTTAATTCATCGA---TTAGTGGAAACTGGCAGGATTTAA 129
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      23 ---LeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla 41
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      130 GATTTTAAAAATGAACCTCAAGAACGAG---TATAGAGTAAATTACTTATGATGTCAGAGT 186
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      42 SerGlyThrAspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuPro 61
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      187 CATGGAAAACTTTCACGA-----ACAGAATCATATGAATTTAAAGATCATGTTTGAAGT 240
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      62 LeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSer 81
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      241 TTAATGATTTAATGGGAGCATTAATAATCGATTCGCACATATTTTA---GGACATGAT 297
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      82 LeuGlyGlyMetAsnLeuLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAla 101
          ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db      298 ATGGGGGCATCATTCGCAGTGATTTACTGAAAAAATATCAATATAAGATGATTACATTG 357
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      102 ValPheLeuAlaAlaPheMetProAspSerValHisAsnSerPheValLeuGluGln 121
          ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db      358 ACAATTGTTTCGGCCAAAGTGAAGACATTCGAAATGGTTTCAACAAATTAATGTTGAT 417
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      122 TyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySer 141
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      418 TACCAAGAGAATTAGCAGGCTTTAATAAATCTCAGGCAATGATTTAT----- 465
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      142 ProGluGluProLeuThrSerMetPheGlyProLys-----Phe 155
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      466 -----TTATTTCTTAATTTAATAAGAGAAAGATAAGCAATGCAATGGTAT 513
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      156 LeuAlaHisLysLeuTyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeu 175
          :::: ||||| ||||| ||||| ||||| ||||| |||||

```

```
QY 104 LeuAlaAlaPheMetProAspSerValHisAenSerSerPheValLeuGluGlnTyrAen 123
   :::::|||||:::|||||:::
Db 355 ATTGAGCCTTTCTC-----CGCAACGGTGGGAGCAAGATGATGCTCTACGTC 402

QY 124 GluArgThrPro--AlaGluAenTrpLeuAapThrGlnPheLeuProTyrGlySerPro 142
   :::::|||||:::|||||:::
Db 403 TTACGTATGGCACTCTACGCTTGGGGGCCCACTGTATGGCGGAGTTATGCCCTCTGGC 462

QY 143 GluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGln 162
   |||||:::|||||:::
Db 463 CTCTGGCCT-----GGACTGGCGAGAGCGCGCTGCCGAGCGCG----- 504

QY 163 LeuCysSerProGluAapLeuAlaLeuAaSerSerLeuValArgProSer-----Ser 180
   |||||:::|||||:::
Db 505 -----GCCCGCACAACTGCAATGCTGACCGCACCGGAGAGATGGAGC 546

QY 181 LeuPheMetGluAapLeuSerLysAlaLysTyrPheThrAapGluArgPhe---GlySer 199
   |||||:::|||||:::
Db 547 GCCTTTCGCTCCACCGCTGGCGCGCTGCGACCAAGTGTGTGAGCCCTACCTTGGAAAG 606

QY 200 ValLysArgValTyrIleValCysThrGluAapLys-----GlyIleProGlu 215
   |||||:::|||||:::
Db 607 GTGAGGTCACGAGCTGGTGTGTTATGGCGACTCCGACCGGACTGCTCAGCCTTTG 666

QY 216 GluPheGlnArgTrpGlnIleAapAenIleGlyValThrGluAlaIleLysGly 235
   |||||:::|||||:::
Db 667 GAGGAGGCCAAGTGGTGTGCATCCAACTTCGAAGATGCTGAGACATATTGCCGTGGCTGT 726

QY 236 AlaAapHisMetAlaMetLeuCysGluProGlnLysLeuCysAlaSerLeuGluIle 255
   |||||:::|||||:::
Db 727 GCTGGTCAGCTCCTCAATTTGAGAATCTGACGTTGCGACCGAGCGGTCAATTCATTC 786

QY 256 AlaHisLysTyrAen 260
   |||||:::
Db 787 TTGAGCAAGCTAAAT 801
```

## RESULT 5

```
US-09-710-279-3711/c
; Sequence 3711, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3711
; LENGTH: 4331
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-3711
```

```
Alignment Scores:
Pred. No.: 7,39e-05 Length: 4331
Score: 123.00 Matches: 53
Percent Similarity: 44.39% Conservative: 34
Best Local Similarity: 27.04% Mismatches: 85
Query Match: 8.99% Indels: 24
DB: 4 Gaps: 9
```

US-10-780-002-2 (1-260) x US-09-710-279-3711 (1-4331)

```
QY 3 GluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyrLys 22
   |||||:::|||||:::
Db 610 GAAGGATCCCCATCATTTTAATCATGGA---TTAGATGGAAACTTGGCAGGATTTAA 554
```

```
QY 23 ---LeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAapLeuAlaAla 41
   |||||:::|||||:::
Db 553 GATTATAAAAAATGAACCTCAAGAACGAG---TATAGAGTAATATTCTTATGATGTGAGAGGT 497

QY 42 SerGlyThrAapLeuAaGlyLysIleGluGluLeuArgThrLeuTyrAapTyrThrLeuPro 61
   |||||:::|||||:::
Db 496 CATGGAATAATCTTCACGA-----ACAGATCATATGAATTAAGAGATCATGTTGAAGAT 443

QY 62 LeuMetGluLeuMetGluSerLeuSerAlaAapGluLysValIleLeuValGlyHisSer 81
   |||||:::|||||:::
Db 442 TTAATGATTAAATGGGAGCATTAATAATATCGATTCTGCACATATTTTA---GGACATGAT 386

QY 82 LeuGlyGlyMetAenLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAla 101
   :::::|||||:::|||||:::
Db 385 ATGGGGGGCATCATTTGGCAGTGAATTTACTGAAAAAATATCAATATAAAGTGAATACATTG 326

QY 102 ValPheLeuAlaAlaPheMetProAapSerValHisAenSerSerPheValLeuGluGln 121
   :::::|||||:::|||||:::
Db 325 ACAATTGTTTCGGCCAAAAGTGAAGACATTGCCAATGGTTTCAACAAATTAATGTTGAT 266

QY 122 TyrAenGluArgThrProAlaGluAaenTrpLeuAapThrGlnPheLeuProTyrGlySer 141
   |||||:::|||||:::
Db 265 TACCAAGAAGAATTTAGCAGGCTTTAATAAAATCTCGAGCAATGATTATT----- 218

QY 142 ProGluGluProLeuThrSerMetPhePheGlyProLys-----Phe 155
   |||||:::|||||:::
Db 217 -----TTATTCTCTAAATTTTAAAGAGAAAGATAAAGCAATGAAATGGTAT 170

QY 156 LeuAlaHisLysLeuTyrGlnLeuCysSerProGluAapLeuAlaLeuAlaSerSerLeu 175
   :::::|||||:::|||||:::
Db 169 CAAGCCAAAATTTATACATAGACCAACTCCGAGAGAAAGTGCANATGCA----- 119

QY 176 ValArgProSerSerLeuPheMetGluAapLeuSerLysAlaLysTyr 191
   |||||:::|||||:::
Db 118 GTACGT---GCATTGCTTAATATTAAAGATTAACTCGTGTTCATCAT 74

RESULT 6
US-08-106-761-1
; Sequence 1, Application US/08106761
; Patent No. 5445956
; GENERAL INFORMATION:
; APPLICANT: HAMMOCK, Bruce D.
; APPLICANT: GRANT, David F.
; APPLICANT: BEETHAM, Jeffrey K.
; TITLE OF INVENTION: RECOMBINANT SOLUBLE EPOXIDE HYDROLASE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND KHOURIE and CREW
; STREET: Steuart Street Tower, 20th Floor, One Market
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/106,761
; FILING DATE: 19930813
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 2307E-445
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2101 base pairs
```

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 42..1703
US-08-106-761-1

Alignment Scores:
Pred. No.: 4,91e-05 Length: 2101
Score: 120.50 Matches: 59
Percent Similarity: 38.16% Conservative: 28
Best Local Similarity: 25.88% Mismatches: 80
Query Match: 8.81% Indels: 61
DB: 1 Gaps: 11

US-10-780-002-2 (1-260) x US-08-106-761-1 (1-2101)
QY 6 HisPheVal-----LeuValHisGlyAlaCysHisGlyGly 17
Db 792 CATTTTGTGAGCTGGCTGGCTGTGTGCTCTGCGCATGGATTTCGCGAGATTGG 851
QY 18 TrpSerTrpTyrLysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeu 37
Db 852 TATTCCTTGAGGTTACGAGATCCCTGCTCTGGCCGAGGAGTTACCGGTCTCTAGCTATG 911
QY 38 AspLeuAlaAlaSerGlyThrAspLeuArgLysIleGluGluLeuArgThrLeuTyrAsp 57
Db 912 GACATGAAGGCTATGAGAGTATCTGCTCTCCCTCCGAA-----ATAGAGAA 959
QY 58 TyrThrLeu-----ProLeuMetGluLeuMetGluSerLeuSerAlaAsp 72
Db 960 TATTGCTGAAGTATTATGTAAGGAGATGTAACCTTCTGTGATAAACTGGGCGCTCTCT 1019
QY 73 GluLysValIleLeuValGlyHisSerLeuGlyMetAsnLeuGlyLeuAlaMetGlu 92
Db 1020 CAA---GCAGTGTTCATTGGCCATGACTGGGGTGGCATCTGCTGTGTGG----- 1064
QY 93 LysTyrProGlnLysIleTyrAlaAlaValPheLeuAlaAlaPheMetProAspSerVal 112
Db 1065 -----TACATGGCTCTTC----- 1079
QY 113 HisAsnSerSerPheValLeuGluGlnTyrAsnGluArgThrProAlaGluAsnTrpLeu 132
Db 1080 -----TACCCCGAGAGAGTGGGCGGTGGCCAGTTTG 1112
QY 133 AspThrGlnPheLeuProTyrGlySerProGluGluProLeuThrSerMetPheGly 152
Db 1113 AATACCTCCCTTCATACAGCAATCCCAACATGCTCCCTTTTGAGAGTATCAAGCCAAC 1172
QY 153 ProLysPheLeuAlaHisLysLeuTyrGlnLeuCysSerProGluAspLeuAlaLeuAla 172
Db 1173 CCAGTATTT-----GATTACAGCTCTACTTCCAGAACCCAGGATGGCTGAG 1220
QY 173 SerSerLeuValArgProSerSerPheMetGluAspLeuSerLysLysTyrPhe 192
Db 1221 GCTGAACCTGGACAGAACCTAGTGGAGCTTTTCAAAAGCCTCTTCAGAGCA----- 1271
QY 193 ThrAspGluArgPheGlySerValLysArgValTyrIleValCysThrGluAspLysGly 212
Db 1272 AGCGATGAG-----AGTGTGTTTATCCATGATGATAAAGTCTGT---GAAGCGGAGGA 1319
QY 213 Ile-----ProGluGlu 216
Db 1320 CTTTTTGTAAATAGCCAGAGAG 1343

RESULT 7
US-09-902-540-4129
; Sequence 4129, Application US/09902540
```

```
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 4129
; LENGTH: 1107
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-4129

Alignment Scores:
Pred. No.: 5,01e-05 Length: 1107
Score: 117.00 Matches: 65
Percent Similarity: 41.73% Conservative: 41
Best Local Similarity: 25.59% Mismatches: 106
Query Match: 8.55% Indels: 42
DB: 4 Gaps: 14

US-10-780-002-2 (1-260) x US-09-902-540-4129 (1-1107)
QY 3 GluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyrLys 22
Db 154 GACGCGAAGCGCGTCTTTCATCCACGCGCTGGTTCGTACCTGAAGTCTTGGCGGGCG 213
QY 23 LysLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAsp 42
Db 214 CAGCTGGACGCTTCCAGAACGAGGCTACCGCGTCATCGCGGTGGACTGCCCGGCTAC 273
QY 43 GlyThrAspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeu 62
Db 274 GGC-----AATCCGACAAAGCCCGCACCC-----TTCCCGTACCATGGAGGCC 318
QY 63 Met-----GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeu 77
Db 319 ATGGCGGACGCGGTGCTGAGTTGGTGACGCGCTGGGCTGGAC---AAGCCCGTGTCTC 375
QY 78 ValGlyHisSerLeuGlyGlyMet---AsnLeuGlyLeuAlaMetGluLysTyrProGln 96
Db 376 GCGGTCATCTCATGGCGGGGACGACGTCGTCTCTTCCGCAATC---CGTACCCCGGAG 432
QY 97 LysIleTyrAlaAlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSer 116
Db 433 TCGCTGACGCGGCTGCTGCTGCGCTGCGCC-----GCCGCGC 468
QY 117 PheValLeuGluGlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPhe 136
Db 459 TTCGAGAAGTTTCAGCTGGCGGAGAAAGAGTGTTCGCGCGGCGCATGATGATCCGAGTTC 528
QY 137 LeuProTyrGlySerProGluGluProLeu-----ThrSerMet 149
Db 529 ATCAAG---TCCGCGCGCGGAGGCCAGCATCTGGGGCAGCGTGGCGCACGCGCAACTTCATG 585
QY 150 PhePheGlyPro-----LysPheLeuAlaHisLysLeuTyrGlnLeuCys---SerPro 166
Db 586 CACTGGCGCGCGAGTTGAGTGGCTGATGAGGAGCGCGCTGGCGCTGGCGAAGTCGCCC 645
QY 167 GluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMetGluAspLeu 186
Db 646 GAGTTCGACGCTACGCTACGCCAACGTGCGCAGC-----GTGCGAGCGCTTC 693
QY 187 SerLysAlaLysTyrPheThrAspGluArgPheGly---SerValLysArgValTyrIle 205
Db 694 TCCCAACACGACTTCGTGGCGGCAACACCTGCACCGCGTCCACCGTCCCGCATCATC 753
```







QY 142 ProGluGluProLeuThrSerMetPheGlyPro-Lys-----Ph 155  
Db 1074 -----TTATCTCTAAATTAATAAGAGAAAGATAAGCAATGAATGGGTA 1121  
QY 155 eLeuAlaHisLysLeuTyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLe 175  
Db 1122 TCAAGGCCAAATTAATATACATAGACCAACTCCGGAAGAAAGTGCAATTGCA----- 1173  
QY 175 uValArgProSerSerLeuPheMetGluAspLeu----- 186  
Db 1174 -GTACGT---GCATGCTTAATATTAAGATTAACTCGTTCATCATATAATGTGCCAT 1229  
QY 187 -----SerLysAlaLysTyrPh 192  
Db 1230 ACCTACTTTAATGTGAATGGTAAGTATGACCACTCATACAAAATAAAGTCATTATGA 1289  
QY 192 eThrAspGluArgPheGlySerValLysArgValTyrIleValCysThrGluAspLysG1 212  
Db 1290 TATGATCAATATTATGATCAAGTTACAAAATTGTATT-----GATAATTCAGG 1340  
QY 212 Y-----IleProGluGluPheGlnArgTyrGlnIleAsp 223  
Db 1341 ACATGCACCACATATCGAGGAACCAAGAAAATTCCTGAAACTCTACTTAGAT 1392

## RESULT 11

US-08-921-177-6  
; Sequence 6, Application US/08921177  
; Patent No. 5798211  
; GENERAL INFORMATION:  
; APPLICANT: Ohno, Tsuneya  
; APPLICANT: Matsuhisa, Akio  
; APPLICANT: Uehara, Hirotugu  
; APPLICANT: Eda, Soji  
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/921,177  
; FILING DATE: 29-AUG-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,577  
; FILING DATE: 27-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 19036/32420  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8654 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
; ORIGINAL SOURCE:  
; ORGANISM: Staphylococcus epidermidis  
; STRAIN: Clinical isolate SE-22  
; US-08-921-177-6

Alignment Scores:  
Pred. No.: 0.0033 Length: 8654  
Score: 114.00 Matches: 62  
Percent Similarity: 40.31% Conservative: 42  
Best Local Similarity: 24.03% Mismatches: 97  
Query Match: 8.33% Indels: 57  
DB: 1 Gaps: 12  
US-10-780-002-2 (1-260) x US-08-921-177-6 (1-8654)  
QY 3 GluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTyrTyrLys 22  
Db 681 GAAGGCATCCCATCATTTTAATTCATGGA---TTAGATGCAAACTGGCAGGATTAAA 737  
QY 23 ---LeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla 41  
Db 738 GATTTAAAAAATGAATCAAGAGCAG---TATAGAGTAATTACTTATGATGTGAGAGGT 794  
QY 42 SerGlyThrAspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuPro 61  
Db 795 CATGAAAATCTTCACGA-----ACAGATCATATGAATTAAGATCATGTTGAAGAT 848  
QY 62 LeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSer 81  
Db 849 TTAATGATTTAATGGGAGCATTAATAATCGATCTGCACATATTTA---GGACATGAT 905  
QY 82 LeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAla 101  
Db 906 ATGGGGGCATCATTCGAGTGAATTTACTGAAAAATATCAATATATAAAGTGATACATTG 965  
QY 102 ValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGln 121  
Db 966 ACAATTGTTTCGGCCAAAAGTGAAGACATTCGAATGGTTTCAACAAATTAATGTTGAT 1025  
QY 122 TyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySer 141  
Db 1026 TACCAGAGAAGATAGCAGGGCTTTAATAAATCTGAGGCAATGATTATT----- 1073  
QY 142 ProGluGluProLeuThrSerMetPheGlyPro-Lys-----Ph 155  
Db 1074 -----TTATCTCTAAATTAATTAAGAGAAAGATAAAGCAATGAATGGGTA 1121  
QY 155 eLeuAlaHisLysLeuTyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLe 175  
Db 1122 TCAAGGCCAAATTAATATACATAGACCAACTCCGGAAGAAAGTGCAATTGCA----- 1173  
QY 175 uValArgProSerSerLeuPheMetGluAspLeu----- 186  
Db 1174 -GTACGT---GCATGCTTAATATTAAGATTAACTCGTTCATCATATAATGTGCCAT 1229  
QY 187 -----SerLysAlaLysTyrPh 192  
Db 1230 ACCTACTTTAATGTGAATGGTAAGTATGACCACTCATACAAAATAAAGTCATTATGA 1289  
QY 192 eThrAspGluArgPheGlySerValLysArgValTyrIleValCysThrGluAspLysG1 212  
Db 1290 TATGATCAATATTATGATCAAGTTACAAAATTGTATT-----GATAATTCAGG 1340  
QY 212 Y-----IleProGluGluPheGlnArgTyrGlnIleAsp 223  
Db 1341 ACATGCACCACATATCGAGGAACCAAGAAAATTCCTGAAACTCTACTTAGAT 1392

## RESULT 12

US-08-362-577C-6  
; Sequence 6, Application US/08362577C  
; Patent No. 5807673  
; GENERAL INFORMATION:  
; APPLICANT: Ohno, Tsuneya  
; APPLICANT: Matsuhisa, Akio  
; APPLICANT: Uehara, Hirotugu  
; APPLICANT: Eda, Soji  
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease





;  
; STRAIN: Clinical Isolate SE-22  
; US-08-920-828-6

## Alignment Scores:

Pred. No.: 0.0033 Length: 8654  
Score: 114.00 Matches: 62  
Percent Similarity: 40.31% Conservative: 42  
Best Local Similarity: 24.03% Mismatches: 97  
Query Match: 8.33% Indels: 57  
DB: 2 Gaps: 12

US-10-780-002-2 (1-260) x US-08-920-828-6 (1-8654)

```
QY 3 GluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyrLys 22
Db 681 GAAGGCATCCCCATCATTTTAAATTCATGGA---TTAGATGGAAACTTGGCAGGATTAAA 737
QY 23 ---LeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla 41
Db 738 GATTTAAAAAATGAATCAAGAACGAG---TATAGAGTAATTACTTATGATGTCAGAGGT 794
QY 42 SerGlyThrAspLeuArgLysIleGluLeuArgThrLeuTyrAspTyrThrLeuPro 61
Db 795 CATGGAAATCTTCACGA-----ACAGAATCATATGAATTAAGAGATCATGTTCAAGAT 848
QY 62 LeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSer 81
Db 849 TTAATATGATTTAATGGGACATTAATATCGATTCGCACATATTTTA---GGACATGAT 905
QY 82 LeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAla 101
Db 906 ATGGGGGGCATCATTCGCGAGTGAAATTACTGAAATAATATCAATATAAGTGAATTACATG 965
QY 102 ValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGln 121
Db 966 ACAATGTTTCGCCCAAAAGTGAACACATTCGAATGTTTCAACAAATTAATGTTGAT 1025
QY 122 TyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySer 141
Db 1026 TACCAAGAAGAATTAGCAGCGCTTTAATAATCTGAGGCAATGATTATT----- 1073
QY 142 ProGluGluProLeuThrSerMetPhePheGlyPro-Lys-----Ph 155
Db 1074 -----TTATCTCTAAATTTTAAAGAGAAGATAAGCAATGAATGGGTA 1121
QY 155 eLeuAlaHisLysLeuTyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSert 175
Db 1122 TCAAAAGCCAAAATATATACAAATAGACCAACTCGGAAGAAAGTGCAATTGCA----- 1173
QY 175 uValArgProSerSerLeuPheMetGluAspLeu----- 186
Db 1174 -GTACGT---GCATTGCTTAAATATTAAAGATTAACTCGTGTTCATCATTAATGTGTCCAT 1229
QY 187 -----SerLysAlaLysTyrPh 192
Db 1230 ACTACTTTAATGTGAATGGTAAGTATGACCACCTACATAAATAAAGTCATATGA 1289
QY 192 eThrAspGluArgPheGlySerValLysArgValTyrIleValCysThrGluAspLysG 212
Db 1290 TATGATCATATATGATCAAGTTACAAAATTTGATTT-----GATAATTCAGG 1340
QY 212 Y-----IleProGluGluPheGlnArgTrpGlnIleAsp 223
Db 1341 ACATGCACCACCATATCGAGGAACCCAGAAAATTTCTGAAACTCTACTTAGAT 1392
```

## RESULT 14

US-09-252-991A-9351/c  
; Sequence 9351, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

;  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 9351  
; LENGTH: 657  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
; US-09-252-991A-9351

Alignment Scores:  
Pred. No.: 7.28-05 Length: 657  
Score: 113.00 Matches: 32  
Percent Similarity: 46.15% Conservative: 22  
Best Local Similarity: 27.35% Mismatches: 45  
Query Match: 8.28% Indels: 18  
DB: 4 Gaps: 4

US-10-780-002-2 (1-260) x US-09-252-991A-9351 (1-657)

```
QY 3 GluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyrLys 22
Db 610 GACGCAAGCCGGTATCGCCCTGCATGGCTGGCTGGACAACGCCATGAGTTCTCGCGC 551
QY 23 LeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAlaSer 42
Db 550 CTGGCGCGGAAGCTC-----GCCGGGCTGGTATCGTCGCCCTCGACTTCGCCGGCCAC 497
QY 43 GlyThrAspLeuArgLysIleGluGluLeuArg---ThrLeuTyrAspTyrThrLeuPro 61
Db 496 GGGCATTCGCGGCATCTGTGCGCGGCGGCCAGCTACCTGCTCTGGGACTACGCGCTGGAC 437
QY 62 LeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSer 81
Db 436 GTGCTGATGGTGGCGGAGCAACTGGGCTGG---GAGCGTTTTTCTCTGTTGGGCACTCG 380
QY 82 LeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAla 101
Db 379 ATGGCGGCCATC-----GTCTCG 362
QY 102 ValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheVal 118
Db 361 GTGCTGCTCGCGGGGCTTGCCTCCGAGCATCGAGCGGTGGCGTGTGATC 311
```

## RESULT 15

US-09-252-991A-9422  
; Sequence 9422, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 9422  
; LENGTH: 855  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
; US-09-252-991A-9422

Alignment Scores:  
Pred. No.: 0.00011 Length: 855  
Score: 113.00 Matches: 32

Percent Similarity: 46.15%  
 Best Local Similarity: 27.35%  
 Query Match: 8.26%  
 DB: 4  
 Conservative: 22  
 Mismatches: 45  
 Indels: 18  
 Gaps: 4

US-10-780-002-2 (1-260) x US-09-252-991A-9422 (1-855)

```

QY 3 GluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyrLys 22
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 GACGGCAAGCCGGTGTATGCCCTGATGCTGGCTGGACACGCCCATGACTTCTCGCGC 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 23 LeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAsp 42
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 CTGGCGCCGAAGTC-----GCCGGCTGCGTATGCTGCGCTCGACTTGGCGGCCAC 235
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 43 GlyThrAspLeuArgLysIleGluGluLeuArg--ThrLeuTyrAspTyrThrLeuPro 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 236 GGGCATTTCCGCGCATCGTCCGAGGGCGCCAGCTACTGCTCTGGGACTACGCGCTGGAC 295
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 62 LeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSer 81
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 296 GTGCTGATGGTGGCCGAGCAACTGGGCTGG---GACCGTTTTTCCCTGTGGGGCACTCG 352
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 82 LeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAla 101
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 353 ATGGGCGCCATC-----GTCTCG 370
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 102 ValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheVal 118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 371 GTGCTGCTCGCGGGGCTTGGCCGAGCGCATCGAGCGGCTGGCGTTGATC 421
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    
```

Search completed: September 12, 2005, 22:51:02  
 Job time : 202 secs

**This Page Blank (uspto)**



PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

|                       |                 |   |           |             |
|-----------------------|-----------------|---|-----------|-------------|
| Query Match           | 24.0%;          | Score 218.8;  | DB 3;     | Length 792; |
| Best Local Similarity | 56.6%;          | Pred. NO. 4.1e-60;  |           |             |
| Matches 447;          | Conservative 0; | Mismatches 337;   | Indels 6; | Gaps 2;     |
| QY                    | 55              | GAGACATATGGAGAGAAACATCATTGCTGTAGTTCACAGCGTTATCATGGAGCCCTG     | 114       |             |
| Db                    | 3               | GAGTGAGGAGGAGAGAGCAACACGTCGTTCTAGTACATGGTCTTGCCATGGCGCTG      | 62        |             |
| QY                    | 115             | GATCTGGTACAGCTCAAGCCCTCTTGAATCAGCCGCGCCACCGGTTACTGTGTCGA      | 174       |             |
| Db                    | 63              | GTGCTGGTCAAGGTTAAGCCGAGCTCGAGGCTTCTGGCCACCGCGTAAACCGCGTAGA    | 122       |             |
| QY                    | 175             | ACTCGCGCTCCGGGATCGAC---CCACGACCAATCCAGGCGTTGAAACCGTCCAGCA     | 231       |             |
| Db                    | 123             | TCTAGCTGCTCCGGTATAGACATGACACGAGTCAATCAGATATATCCACATCGCAACA    | 182       |             |
| QY                    | 232             | ATACTCCAAACCGTTGATCGAAACCCCTCAAACTCTTCCAGAGAACGAAAGAGTAACTTCT | 291       |             |
| Db                    | 183             | ATACTCAGAGCCATTGATGACGCTAATGACCTCACTACCAGATGATGAGAAGGTTGTGCT  | 242       |             |
| QY                    | 292             | GTTTGAATTCAGCTTCGGAGGCATCAACATCGCTCTCGCGCGGACATATTTCCGCGAA    | 351       |             |
| Db                    | 243             | TGTTGTCATAGCTTAGGAGGTTTGAGTTAGCTATGGCCATGGATATGTTTCCGACCAA    | 302       |             |
| QY                    | 352             | GATTAAGGTTCTGCTCTCAACGCTTCTTCCCGGACACAAACCCAGTGGCTTCTCA       | 411       |             |
| Db                    | 303             | AACTCTGTTCTGCTTCTTGTGACTGCTATGATGCCAGACACCAACATCTCACATCTT     | 362       |             |
| QY                    | 412             | CGTCTTGGCAAGTATATGGAGATGCTCGAGGTTTGGAGATTTGGA---GTTTTCATP     | 468       |             |
| Db                    | 363             | CGTATGGGTAAGCTTAAGAAAGAAACTTCAAGAGGAATGTTAGACACCGTGTATAC      | 422       |             |
| QY                    | 469             | TCATGAAACAGAAATGGGACGATGATTTATTTGAAGATGGGACCAAAATTCATGAAGC    | 528       |             |
| Db                    | 423             | GAGCGAGAAACCTGATTTTCTAGCGAGTTTGGATTTTGGACCAAGATTCATGGCCAA     | 482       |             |
| QY                    | 529             | AGCTCTTTACCAAAATGTTCCCATAGAGATTACGAGCTGGCAAAATGTTGCATAGCA     | 588       |             |
| Db                    | 483             | GAACTTGATCAGTTGCTTCCAGTCCAGATCTTGAAATTTGGCAAAATGTTGTGAGGC     | 542       |             |
| QY                    | 589             | AGGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTTCAGGAGGAGATATGG     | 648       |             |
| Db                    | 543             | AAACCCATTGATTTAAGAAAGATATGCGAGAGAGAAAGCTTCAGTGAGGAGGATACGG    | 602       |             |
| QY                    | 649             | TTCCGTGCAACGATTTAGTATGATGATGATGAGACAAAGCCATCCCTGCGATTTCA      | 708       |             |
| Db                    | 603             | ATCCGTTACACGATATTTATTTGATGCGGAAAGGATCTTGTGTACCCGGAAGATTACCA   | 662       |             |
| QY                    | 709             | TGTTGGATGATTTGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGGATCA   | 768       |             |
| Db                    | 663             | GCATCGATGATCAGCACTTTCCCAAGAGATGATGAGATCAAGAGCGAGATCA          | 722       |             |
| QY                    | 769             | CATGGTATGCTCTCAAAACCCCAAAACTCTTTGACTCTCTCTGCTATTTGCCACGA      | 828       |             |
| Db                    | 723             | TATGCCAATGTTCTCCAAGCTCAACACTATGTGCTCTCTCTTCTTGGAGATTGCAATAA   | 782       |             |
| QY                    | 829             | TTATATGTAA  | 838       |             |
| Db                    | 783             | ATATGCCCTAA   | 792       |             |

RESULT 13

ABZ12412

ID ABZ12412 standard; DNA; 792 BP.

XX

AC ABZ12412;

XX

DT 21-JAN-2003 (first entry)

XX

DE Arabidopsis thaliana stress regulated gene SEQ ID NO 217.

XX

KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX Arabidopsis thaliana.  
OS WO200216655-A2.  
XX 28-FEB-2002.  
XX 24-AUG-2001; 2001WO-US026685.  
XX 24-AUG-2000; 2000US-0227866P.  
PR 26-JAN-2001; 2001US-0264647P.  
PR 22-JUN-2001; 2001US-0300111P.  
XX (SCRI ) SCRIPPS RES INST.  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX Harper JF, Kreps J, Wang X, Zhu T;  
XX WPI; 2002-304127/34.  
XX Identifying a stress condition to which a plant cell has been exposed and  
producing plants with increased tolerance to these abiotic stresses.  
PS Claim 144; SEQ ID NO 217; 577pp + Sequence Listing; English.  
XX The invention relates to identifying a stress condition to which a plant  
cell has been exposed, comprising: (a) contacting nucleic acid  
representative of expressed polynucleotides in the plant cell with an  
array or probes representative of the plant cell genome; and (b)  
detecting a profile of expressed polynucleotides in the plant cell  
characteristic of a stress response. The method is useful in the  
production of transgenic plants, cells and seeds and in producing plants  
with increased tolerance to abiotic stress. The present sequence is that  
of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
in methods of the invention. Note: The sequence data for this patent is  
not represented in the printed specification but is based on sequence  
information supplied to Derwent by the European Patent Office  
XX Sequence 792 BP; 230 A; 167 C; 192 G; 203 T; 0 U; 0 Other;  
SQ

|                       |                 |   |           |             |
|-----------------------|-----------------|---|-----------|-------------|
| Query Match           | 24.0%;          | Score 218.8;  | DB 6;     | Length 792; |
| Best Local Similarity | 56.6%;          | Pred. NO. 4.1e-60;  |           |             |
| Matches 447;          | Conservative 0; | Mismatches 337;   | Indels 6; | Gaps 2;     |
| QY                    | 55              | GAGACATATGAGAGAAACATCATTGCTGTAGTTCACACGCTTATCATGGAGCCCTG    | 114       |             |
| Db                    | 3               | GAGTGAGGAGAGAGAGCAACACGTCGTTCTAGTACATGGTCTTGCCATGGCGCTG     | 62        |             |
| QY                    | 115             | GATCTGGTACAGCTCAAGCCCTCTTGAATCAGCCGCGCCACCGCTTACTGTGTCGA    | 174       |             |
| Db                    | 63              | GTGCTGGTACAGGTTAAGCCGAGCTCGAGGCTTCTGGCCACCGGTAACCGCGTAGA    | 122       |             |
| QY                    | 175             | ACTCGCGCTCCGGGATCGAC---CCACGACCAATCCAGGCGGTTGAAACCGTCCAGCA  | 231       |             |
| Db                    | 123             | TCTAGCTGCTCCGGTATAGACATGACAGGTCATACAGATATATCCACATGGCAACA    | 182       |             |
| QY                    | 232             | ATACTCCAAACCGTTGATCGAAACCTCAAACTCTTCCAGAGAACGAGGTAATCTCT    | 291       |             |
| Db                    | 183             | ATACTCAGAGCCATTGATGACGCTTAACTACCTACACATGATGAGAAGTTGTGCT     | 242       |             |
| QY                    | 292             | GTTTGGATTCAGCTTCGGAGGCATCAACATCGCTCTCGCCGCGGACATATTTCCGCGAA | 351       |             |
| Db                    | 243             | TGTTGGTTCATAGCTTAGGAGGTTTGAGTTAGTATGGCCATGGATATGTTCCGACCA   | 302       |             |
| QY                    | 352             | GATTAAGGTTCTTGTGTTCTCAACGCTTCTTGGCCGACACCAACCCAGTGGCTTCTCA  | 411       |             |
| Db                    | 303             | AATCTCTGTTCTGCTTTGTGACTGCTATGATGCGACAGACCAACCACTCACCATCTT   | 362       |             |
| QY                    | 412             | CGTCTCGACAGATATATGGAGATGCCCTGGAGTTTGGGAGATTGTGA---GTTTTCATC | 468       |             |
| Db                    | 363             | CGTATGGGTAAGCTTAAGAAAGAAACTTCCAGAGAGGAATGTTTAGACACCGTGTATAC | 422       |             |
| QY                    | 469             | TCATGAAACAGAAATGGGACGATGAGTTTATTTGAAGATGGGACCAAAATTCATGAAGC | 528       |             |

Db 423 GAGCGAGAAACCTGATTTCTTAGCGAGTTTGGATTTTGGACAGAAATTCATGGCCAA 482  
QY 529 ACCTCTTTACCAAAATGTGCCATAGAGGATTACGAGCTGGCCAAAATGTTGCATAGGCA 588  
Db 483 GAACTTGATCAGTTGTCTCCAGTCCAGATCTTGAATGGCGAAATGTTGGTCAGGGC 542  
QY 589 AGGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTACGAGGAGGATATGG 648  
Db 543 AAACCCATTGATTAAAGAAAGATATGGCAGAGAGAAAGCTTCAGTAGGAGGATACGG 602  
QY 649 TTCGGTGAACGAGTTTACGTAATCAGTAGTGAAGACAAAGCCATCCCTGCGATTTCA 708  
Db 603 ATCCGTTACAGCTATATTATTGTTATGCGGAAAGGATCTTGTGCACCGAAGATTACCA 662  
QY 709 TCGTTGGGATGATTGATAATTTCAACGCTCTCGAAAGCTACAGAGATCGATGCGGAGATCA 768  
Db 663 GCGATCGATGATCAGCAACTTTCCCCCAAGAGAGTAAATGGAGATCAAGAGCGAGATCA 722  
QY 769 CATGGTGTGCTCTCCAAACCCCAAAACTCTTTGCACTCTCTCTGCTATTTGCCACCGA 828  
Db 723 TATGCCAATGTTCTCCAGCCTCAACAACATATGTGCTCTCTCTTCTTGGAGATTGCAAAATA 782  
QY 829 TTATATGTA 838  
Db 783 ATATGCTTAA 792

## RESULT 14

ID ADN74620 standard; cDNA; 792 BP.  
XX AC ADN74620;  
XX 15-JUL-2004 (firet entry)  
XX Thale cress cDNA repressed in E2Fa/Dpa expressing plants SeqID 2515.  
KW gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;  
KW growth regulator; animal feed product; thale cress;  
KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.  
XX Arabidopsis thaliana.  
XX WO2004035798-A2.  
XX 29-APR-2004.  
XX 20-OCT-2003; 2003WO-BP011658.  
XX 18-OCT-2002; 2002EP-00079408.  
XX (CROP-) CROPDISEGN NV.  
XX Inze D, De Veylder L, Vlieghe K;  
XX WPI; 2004-348466/32.  
XX P-PSDB; ADN74621.

PT Altering plant characteristics, useful for producing plants for enzyme or  
PT pharmaceutical production comprises modifying in a plant, expression of  
PT one or more nucleic acids and/or modifying level or activity of one or  
PT more proteins.

PS Claim 1; SEQ ID NO 2515; 134pp; English.

XX This invention relates to a novel method for altering one or more plant  
XX characteristics. Specifically, it refers to identifying genes that are up  
XX - or down-regulated in transgenic plants overexpressing the heterodimeric  
XX E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to  
XX alter plant characteristics accordingly. The present invention describes  
XX generating transgenic plants for the production of growth regulators,  
XX enzymes, therapeutics, pharmaceuticals and animal feed products, where

CC the altered plant characteristics are selected from increased yield or  
CC biomass, enhanced survival capacity, stress tolerance, plant architecture  
CC or physiology, altered endoreplication, biochemistry, signal  
CC transduction, storage lipid mobilisation and/or altered photosynthesis,  
CC each relative to the corresponding wild type plants. Accordingly, these  
CC sequences can also be useful as positive or negative selectable markers  
CC during transformation of cells or tissues. The identified genes play a  
CC role in a variety of biological processes such as DNA replication, cell  
CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as  
CC transcription factors. This polynucleotide sequence is thale cress cDNA  
CC repressed 1.3 fold or more in plants overexpressing the E2Fa/Dpa  
CC transcription factor, given in an exemplification of the invention.

XX Sequence 792 BP; 230 A; 167 C; 192 G; 203 T; 0 U; 0 Other;

QY Query Match 24.0%; Score 218.8; DB 12; Length 792;  
Best Local Similarity 56.6%; Pred. No. 4.1e-60;  
Matches 447; Conservative 0; Mismatches 337; Indels 6; Gaps 2;

QY 55 GAGACATATGAGAGGAAACATCATCTTGTGTTAGTTTACAAACCGTTATCATGAGCGCTG 114  
Db 3 GAGTGAGGAGGAGGAGCAACACGTCGTTCTAGTACATGGTGTCTGCCATGCGCGCTG 62  
QY 115 GATCTGTGTACAGCTCAAGCCCTCTTGAATCAGCGCGCCACCGCTTACTGTGTGCA 174  
Db 63 GTGCTGGTACAAAGGTTAAGCGCAGCTCGAGGCTTCTGGCCACCGCTTAACCGCGGTAGA 122  
QY 175 ACTCGCGCCTCCCGGATCGAC---CCACGACCAATCCAGCCCGTTGAAACCGTCGACGA 231  
Db 123 TCTAGTGCCTCCGTTATAGACATGACAGTCAATCACAGATATATCCACATCGGAACA 182  
QY 232 ATACTCAAAACCGTTGATCGAAACCTCAATCTCTTCCAGAGAACGAGGTAATTTCT 291  
Db 183 ATACTCAGAGCCATTGATGACGTAATGACCTCACTACAGATGATGAGAAAGTTGTGCT 242  
QY 292 GGTGGATTACAGCTTCGAGGCGATCAACATCGCTCTCGCCCGCACATATTTCCGGCGAA 351  
Db 243 TGTGGTCATAGCTTAGGAGGTTTGTAGTTAGTTAGTATGSCCATGATATGTTCCGACCA 302  
QY 352 GATTAAGGTTCTGTGTTCTTCAACGCTTCTTCCCGGACACAAACCGTGTCTTCA 411  
Db 303 AATCTCTGTTCTGTCTTTGTGACTGTCTATGATGCCAGACACCAAACTCACCATCCTT 362  
QY 412 CGTTCTGGAACAAGTATATGGAGATGCTGGAGGTTTGGAGATTTGTGA--GTTTTCATC 468  
Db 363 CGTATGGGATAAGCTAAAGAAAGAACTTCCAGAGAGGAATGGTTAGACACCGTGTTTAC 422  
QY 469 TCATCAACAAAGAAATGGGAGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGGC 528  
Db 423 GAGCGAGAAACCTGATTTTCTAGCGAGTTTGGATTTTGGACAGAAATTCATGGCCAA 482  
QY 529 ACCTCTTTTACCAAAATTTGCCATAGAGGATTACGAGCTGGCCAAAATGTTGCATAGCA 588  
Db 483 GAACTTGATCAGTTGTCTCCAGTCCAGATCTTGAATTTGGCGAAATGTTGGTGAGGGC 542  
QY 589 AGGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTACGAGGAGGATATGG 648  
Db 543 AAACCCATTGATTAAAGAAAGATATGGCAGAGAGAAAGAGCTTCAGTAGGAGGATACGG 602  
QY 649 TTCGGTGAACGAGTTTACGTAATCAGTAGTGAAGACAAAGCCATCCCTCGCATTTTCAT 708  
Db 603 ATCCGTTACAGTATATTTATTTGATGCGGAAAGAGTCTTGTGTACCCGAGGATACCA 662  
QY 709 TCGTTGGATGATGATTAATTTCAACGCTCTCGAAAGTCTTACGAGATCGAGATCGGAGATCA 768  
Db 663 GCGATCGATGATCAGCAACTTTCCCGCAAGAAAGTAAATGGAGATCAAGACGAGATCA 722  
QY 769 CATGGTGTGCTCTCCAAACCCCAAAACTCTTTGCACTCTCTCTGCTATTTGCCACCGA 828  
Db 723 TATGCCAATGTTCTCCAGCCTCAACAACATATGTGCTCTCTTCTTGGAGATTGCAAAATA 782  
QY 829 TTATATGTA 838  
Db 783 ATATGCTTAA 838





**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: September 11, 2005, 22:54:15 ; Search time 3956.73 Seconds  
(without alignments)  
11168.596 Million cell updates/sec

Title: US-10-780-002-36  
Perfect score: 912  
Sequence: 1 atcacgctattctcaacaa.....ttttcgggcaacttctc 912

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_btg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sta.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description         |
|------------|-------|-------------|--------|----|---------------------|
| 1          | 910.4 | 99.8        | 928    | 8  | AY046015 Arabidops  |
| 2          | 882   | 96.7        | 911    | 8  | AY058115 Arabidops  |
| 3          | 808   | 88.6        | 808    | 8  | AY142490 Arabidops  |
| 4          | 777   | 85.2        | 777    | 8  | AY093714 Arabidops  |
| 5          | 628   | 68.9        | 57180  | 8  | ATF18D22 Arabidops  |
| 6          | 341.4 | 37.4        | 856    | 12 | AY202287 Arabidops  |
| 7          | 255.2 | 28.0        | 891    | 8  | AY751530 Catharant  |
| 8          | 246.8 | 27.1        | 1077   | 8  | AF269158 Citrus si  |
| 9          | 241.6 | 26.5        | 792    | 6  | AX506625 Sequence   |
| 10         | 241.6 | 26.5        | 792    | 8  | BT014881 Arabidops  |
| 11         | 241.4 | 26.5        | 935    | 8  | BT002859 Arabidops  |
| 12         | 232.6 | 25.5        | 1079   | 8  | AY485932 Nicotiana  |
| 13         | 221.6 | 24.3        | 1078   | 8  | HEB04042 Hevea bras |
| 14         | 221.6 | 24.3        | 1091   | 6  | AS9586 Sequence 1   |
| 15         | 218.8 | 24.0        | 792    | 6  | CQ806104 Sequence   |
| 16         | 218.8 | 24.0        | 792    | 6  | AX505522 Sequence   |
| 17         | 218.8 | 24.0        | 792    | 8  | BT015031 Arabidops  |
| 18         | 218.8 | 24.0        | 878    | 8  | BT014901 Arabidops  |
| 19         | 214.2 | 23.5        | 792    | 6  | AX506856 Sequence   |

|    |       |      |        |   |                     |
|----|-------|------|--------|---|---------------------|
| 20 | 214.2 | 23.5 | 792    | 8 | AY056068 Arabidops  |
| 21 | 214.2 | 23.5 | 870    | 8 | AY086590 Arabidops  |
| 22 | 214.2 | 23.5 | 932    | 8 | AF361627 Arabidops  |
| 23 | 211   | 23.1 | 1021   | 8 | MEAHXN1 Arabidops   |
| 24 | 210.4 | 23.1 | 961    | 8 | AF178576 Rauvolfia  |
| 25 | 205.4 | 22.5 | 777    | 6 | CQ829357 Sequence   |
| 26 | 204.8 | 22.5 | 1073   | 8 | AK061058 Oryza sat  |
| 27 | 203.4 | 22.3 | 928    | 8 | AK117107 Arabidops  |
| 28 | 199.6 | 21.9 | 792    | 8 | BT010429 Arabidops  |
| 29 | 199.6 | 21.9 | 902    | 8 | AK175772 Arabidops  |
| 30 | 198.6 | 21.8 | 777    | 6 | CQ829358 Sequence   |
| 31 | 196   | 21.5 | 771    | 6 | BT006227 Arabidops  |
| 32 | 194.4 | 21.3 | 771    | 6 | AX506583 Sequence   |
| 33 | 193.4 | 21.2 | 770    | 6 | AX651280 Arabidops  |
| 34 | 185.2 | 20.3 | 944    | 8 | BT010709 Arabidops  |
| 35 | 183.6 | 20.1 | 783    | 8 | BT012414 Arabidops  |
| 36 | 178.6 | 19.6 | 795    | 6 | AX654444 Arabidops  |
| 37 | 176.8 | 19.4 | 924    | 8 | BT004125 Arabidops  |
| 38 | 172.6 | 18.9 | 828    | 8 | AY142031 Arabidops  |
| 39 | 172.6 | 18.9 | 903    | 8 | AY086981 Arabidops  |
| 40 | 172.6 | 18.9 | 904    | 8 | AY074858 Arabidops  |
| 41 | 167.4 | 18.4 | 1087   | 8 | BT012867 Lycopersal |
| 42 | 148.4 | 16.3 | 943    | 8 | AK072603 Oryza sat  |
| 43 | 145.2 | 15.9 | 948    | 8 | OSPIR7BR Arabidops  |
| 44 | 140   | 15.4 | 123360 | 8 | AC003040 Arabidops  |
| 45 | 139.2 | 15.3 | 3948   | 8 | MEAJ3281 Manihot e  |

## ALIGNMENTS

AY046015 Arabidopsis thaliana putative alpha-hydroxynitrile lyase  
(At5g10300) mRNA, complete cds.

AY046015 Arabidopsis thaliana (thale cress)  
FLJ CDNA.  
Accession AY046015.1 GI:15028130  
Version Arabidopsis thaliana  
Keywords Arabidopsis thaliana  
Source Arabidopsis thaliana  
Organism Arabidopsis thaliana  
Reference 1 (bases 1 to 928)  
Authors Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

Arabisopsis Full Length cDNA Clones  
Unpublished  
Title Arabidopsis Full Length cDNA Clones  
Journal Unpublished  
Reference 2 (bases 1 to 928)  
Authors Yamada,K., Liu,S.X., Pham,P.K., Banh,J., Dale,J.M., Gibson,H.A., Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Tang,C.C., Toriumi,M., Yu,G., Yu,S., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Koesema,E., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

Submitted (06-JUL-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  
Riken Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : RIKEN Arabidopsis Full-length cDNA): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the

sequencing and annotation of the RAFL cDNAs: Yamada, K., Liu, S.X., Pham, P.K., Banh, J., Dale, J.M., Gibson, H.A., Goldsmith, A.D., Jiang, P.X., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tracy, S.E., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGE) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGE) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

#### FEATURES

Location/Qualifiers  
 1..928  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="5"  
 /clone="RAFL05-09-N03 (R09923)"  
 /ecotype="Columbia"  
 /note="This clone is in a modified pBluescript vector (lambda ZAP) as a XhoI/SstI insert."  
 1..928  
 /gene="At5g10300"  
 1..61  
 /gene="At5g10300"  
 62..838  
 /gene="At5g10300"  
 /codon\_start=1  
 /evidence=experimental  
 /product="putative alpha-hydroxynitrile lyase"  
 /protein\_id="AAK76689.1"  
 /db\_xref="GI:15028131"  
 /translation="MERKHFFLVHNAVGHAWIYKLPLESLAGHRVTAVELAASGI  
 DRPQAVETVDEYSKPIETLKSIPENEEVILVGSFGGINIALAADIIPAKIKLV  
 FLNAFLPTTHVPSHLVDKMPGLGDCFSHETRTGMLSLKMGKPKMYLQ  
 NCPIEDYELAKMLHQRFFETDLSKKEKFEEGYSVQVYVMSSEDKAIPCFIRW  
 MIDNFNVSKVBIIDGGHMVLSKPKQLFDSLALATDYM"  
 misc\_difference 590  
 /gene="At5g10300"  
 /note="compared to genomic sequence resulting in an amino acid sequence difference"  
 /replace="g"  
 839..928  
 /gene="At5g10300"

#### ORIGIN

Query Match 99.8%; Score 910.4; DB 8; Length 928;  
 Best Local Similarity 99.9%; Pred. No. 6.1e-249;  
 Matches 911; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ATCAGCTATTCTCAACCAAGCAAGAACCTTAACCAAAAGAAAACGCTACGGAGACA 60  
 DB 1 ATCAGCTATTCTCAACCAAGCAAGAACCTTAACCAAAAGAAAACGCTACGGAGACA 60  
 QY 61 TATGAGAGAGAAACATCACTTGGTGTAGTTTCAACACGCTTATCATGAGCGCTGGATCTG 120  
 DB 61 TATGAGAGAGAAACATCACTTGGTGTAGTTTCAACACGCTTATCATGAGCGCTGGATCTG 120  
 QY 121 GTACAAGCTCAAGCCCTCTTGAATCAGCCGCCACCGCTTACTGCTGTCGAACCTGCG 180  
 DB 121 GTACAAGCTCAAGCCCTCTTGAATCAGCCGCCACCGCTTACTGCTGTCGAACCTGCG 180  
 QY 181 CGCTCCGGGATCGACCCAGACCAATCCAGCGCTTGAACCGCTCGAGATATCCAA 240  
 DB 181 CGCTCCGGGATCGACCCAGACCAATCCAGCGCTTGAACCGCTCGAGATATCCAA 240  
 QY 241 ACCGTTGATCGAAACCTCAAAATCTCTTCAGAGAACGAAGAGGTAATTCGTGGTGAATT 300  
 DB 241 ACCGTTGATCGAAACCTCAAAATCTCTTCAGAGAACGAAGAGGTAATTCGTGGTGAATT 300

QY 301 CAGCTTCGGAGGCATCAACATCGCTCTCGCCGCCGACATATTTCCGGCGAAGATTAAAGGT 360  
 DB 301 CAGCTTCGGAGGCATCAACATCGCTCTCGCCGCCGACATATTTCCGGCGAAGATTAAAGGT 360  
 QY 361 TCTTTGTGTTCTCAACGCTTTCTTCCCGGACACAAACCCACGTCCTTCTACGTTCTGGA 420  
 DB 361 TCTTTGTGTTCTCAACGCTTTCTTCCCGGACACAAACCCACGTCCTTCTACGTTCTGGA 420  
 QY 421 CAAGTATATGAGATGCGCTGGAGGTTTGGGAGATTTGAGTTTTCATCTCATGAACAAG 480  
 DB 421 CAAGTATATGAGATGCGCTGGAGGTTTGGGAGATTTGAGTTTTCATCTCATGAACAAG 480  
 QY 481 AAATGGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGGACCGCTTTACCA 540  
 DB 481 AAATGGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGGACCGCTTTACCA 540  
 QY 541 AAATTTGTCCTATAGAGATTACGAGCTGGCAAAATTTGTCATAGGCAAGGTCATTTT 600  
 DB 541 AAATTTGTCCTATAGAGATTACGAGCTGGCAAAATTTGTCATAGGCAAGGTCATTTT 600  
 QY 601 CACAGAGATCTATCAAAAGAAAGATTTAGCGAGGAAGATATGTTTGGTGCACAG 660  
 DB 601 CACAGAGATCTATCAAAAGAAAGATTTAGCGAGGAAGATATGTTTGGTGCACAG 660  
 QY 661 AGTTTACGTAATGAGTAGTGAAGCAAAAGCCATCCCTCGCATTTTCAATTCGTTGGATGAT 720  
 DB 661 AGTTTACGTAATGAGTAGTGAAGCAAAAGCCATCCCTCGCATTTTCAATTCGTTGGATGAT 720  
 QY 721 TGATAAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTGATGCT 780  
 DB 721 TGATAAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTGATGCT 780  
 QY 781 CTCACAAACCCCAAAACCTCTTTGATCTCTCTCTGCTATTCGCCACCGATATATATGTAATA 840  
 DB 781 CTCACAAACCCCAAAACCTCTTTGATCTCTCTCTGCTATTCGCCACCGATATATATGTAATA 840  
 QY 841 ATCTTAAGTCGGTTTACTTTTTTCTCATCGTACTATAAAACAAACCCCTTTTCCGG 900  
 DB 841 ATCTTAAGTCGGTTTACTTTTTTCTCATCGTACTATAAAACAAACCCCTTTTCCGG 900  
 QY 901 GCAACTTTCATC 912  
 DB 901 GCAACTTTCATC 912  
 RESULT 2  
 AY058115 Arabidopsis thaliana AT5g10300/F18D22\_70 mRNA linear PLN 04-NOV-2001  
 LOCUS  
 DEFINITION  
 ACCESSION AY058115  
 VERSION AY058115.1 GI:16648678  
 KEYWORDS  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi  
 1 (bases 1 to 911)  
 Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J.,  
 Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,  
 Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,  
 Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,  
 Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,  
 Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,  
 Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,  
 Davis, R.W., Theologis, A. and Ecker, J.R.  
 Arabidopsis cDNA clones  
 Unpublished  
 2 (bases 1 to 911)  
 Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J.,  
 Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,  
 Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,  
 Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,

Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

## Direct Submission

Submitted (23-AUG-2001) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, P0EC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs (Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Barth, J., Bowser, L., Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

## FEATURES

Location/Qualifiers

1..911

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="5"

/clone="RAFL06-14-F12(R13968)"

/ecotype="Columbia"

1..69

70..846

/note="alpha-hydroxynitrile lyase-like protein"

/cdon\_start=1

/product="AT5G10300/F18D22\_70"

/protein\_id="AAL25532.1"

/db\_xref="GI:16648679"

/translation="MERKHPFLVHNAHYGAMIWYKLPFLLESAGHRVTVAVELASGI

DRPQAVTVDEYSKPLJETLSIQNEEGLVFGGINIALAADIIPAKIKVLV

FLNALPDTHVPSVLDKYMPPGLGDCPSHETRNQWTSLKMGPKMKARLYQ

NCPEDYELAKMLHQKQSFTEDLSEKKEFSEGGYQVORVYVMSSEDAKIPCDPIRW

MIDNFNVSKYBIIDGDHVMLSKPKQKUFDSLISAATDYM"

847..911

## 3' UTR

ORIGIN

Query Match 96.7%; Score 882; DB 8; Length 911;

Best Local Similarity 99.3%; Pred. No. 8.2e-241;

Matches 896; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 ATCACGCTATTCTCAACCAAGCAAGAACTTAACCAAAAGAAACAGCTCTACGGAGACA 60  
 DB |||||  
 DB 10 ATTACGCTATTCTCAACCAAGCAAGAACTTAACCAAAAGAAACAGCTCTACGGAGAC- 68  
 QY 61 TATGGAGAGAAACATCACTTCGTGTAGTTCAACAGCTTATCATGGAGCTTGATCTG 120  
 DB |||||  
 DB 69 TATGGAGAGAAACATCACTTCGTGTAGTTCAACAGCTTATCATGGAGCTTGATCTG 128  
 QY 121 GTACAGCTCAGCCCTCTTGAATCAGCGGCCACCGCTTACTGTCGTAACCTCGC 180  
 DB |||||  
 DB 129 GTACAGCTCAGCCCTCTTGAATCAGCGGCCACCGCTTACTGTCGTAACCTCGC 188  
 QY 181 CGCTCCGGGATCGACCCAGCAACCAATCCAGGCGGTGAAACCGTCGACGAATATCTCAA 240  
 DB |||||  
 DB 189 CGCTCCGGGATCGACCCAGCAACCAATCCAGGCGGTGAAACCGTCGACGAATATCTCAA 248  
 QY 241 ACCGTTGATCGAAACCTCAAAATCTCTTCAGAGACGAGAGGTAATTCGTGGTAT 300  
 DB |||||  
 DB 249 ACCGTTGATCGAAACCTCAAAATCTCTTCAGAGACGAGAGGTAATTCGTGGTAT 308

QY 301 CAGCTTCGAGGAGCATCAACATCGCTCTCGCCGCCGACATATTTCCGCGCAAGATTAAGGT 360  
 DB |||||  
 DB 309 CAGCTTCGAGGAGCATCAACATCGCTCTCGCCGCCGACATATTTCCGCGCAAGATTAAGGT 368  
 QY 361 TCTTGTGTCTCCAAAGCGCTTCTTCCCGACACAAACCAACCGTCTTCTCAGCTTCTGA 420  
 DB |||||  
 DB 369 TCTTGTGTCTCCAAAGCGCTTCTTCCCGACACAAACCAACCGTCTTCTCAGCTTCTGA 428  
 QY 421 CAGTATATATGAGATGCTCGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAACACAG 480  
 DB |||||  
 DB 429 CAGTATATGAGATGCTCGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAACACAG 488  
 QY 481 AAATCGGACGATGAGTTTATTGAAGTGGGACCAAAATTCATGAAGGCAACGCTTTTACCA 540  
 DB |||||  
 DB 489 AAATCGGACGATGAGTTTATTGAAGTGGGACCAAAATTCATGAAGGCAACGCTTTTACCA 548  
 QY 541 AAATGTCCCATAGAGATTACGAGCTGGCAAAATTTGGCATAGGCAAGGTCATTTT 600  
 DB |||||  
 DB 549 AAATGTCCCATAGAGATTACGAGCTGGCAAAATTTGGCATAGGCAAGGTCATTTT 608  
 QY 601 CACAGAGATCTATCAAGAAAGAAAGTTTAGCGGAGGAGGATATGTTCCGTCGCAAGC 660  
 DB |||||  
 DB 609 CACAGAGATCTATCAAGAAAGAAAGTTTAGCGGAGGAGGATATGTTCCGTCGCAAGC 668  
 QY 661 AGTTTACGTAATGAGTAGTGAAGCAAAAGCCATCCCTCGGATTTTCATTCGTTGGATGAT 720  
 DB |||||  
 DB 669 AGTTTACGTAATGAGTAGTGAAGCAAAAGCCATCCCTCGGATTTTCATTCGTTGGATGAT 728  
 QY 721 TGATTAATTTCAACGCTCTCGAAAGTCTACGAGATCGATCGCGAGATCACATCGTGATGCT 780  
 DB |||||  
 DB 729 TGATTAATTTCAACGCTCTCGAAAGTCTACGAGATCGATCGCGAGATCACATCGTGATGCT 788  
 QY 781 CTCGAAACCCCAAAACCTTTTGACTCTCTCTGCTATTGCCACCGATTATATGTAATA 840  
 DB |||||  
 DB 789 CTCGAAACCCCAAAACCTTTTGACTCTCTCTGCTATTGCCACCGATTATATGTAATA 848  
 QY 841 ATCTTAAGTCGCTTTTACTTTTCTCATCGTTACTTAATAAAACAAACCCCTTTTCCGG 900  
 DB |||||  
 DB 849 ATCTTAAGTCGCTTTTACTTTTCTCATCGTTACTTAATAAAACAAACCCCTTTTCCGG 908  
 QY 901 GC 902  
 DB ||  
 DB 909 GC 910  
 QY 901 GC 902  
 DB ||  
 DB 909 GC 910  
 RESULT 3  
 AY142490  
 LOCUS  
 DEFINITION  
 Arabidopsis thaliana putative alpha-hydroxynitrile lyase  
 (AC5G10300) mRNA, complete cds.  
 ACCESSION  
 AY142490  
 VERSION  
 AY142490.1 GI:23296321  
 KEYWORDS  
 FLI CDNA.  
 SOURCE  
 Arabidopsis thaliana (thale cress)  
 ORGANISM  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids;  
 1 (bases 1 to 808)  
 Yamaoka, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W.,  
 Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K.,  
 Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T.,  
 Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G.,  
 Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.  
 Arabidopsis Open Reading Frame (ORF) Clones  
 TITLE  
 Arabidopsis Open Reading Frame (ORF) Clones  
 JOURNAL  
 Unpublished  
 2 (bases 1 to 808)  
 Yamaoka, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W.,  
 Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K.,  
 Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T.,  
 Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G.,  
 Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.



this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (ssp/salk) contributed equally to this work as PIs.

| FEATURES |  | Location/Qualifiers   |   |
|----------|--|---|---|
| source   | 1..777   | /organism="Arabidopsis thaliana"                                  |   |
|          |  | /mol_type="mRNA"  |   |
|          |  | /db_xref="taxon:3702"   |   |
|          |  | /chromosome="5"   |   |
|          |  | /clone="U13968"   |   |
| CDS      |  | /ecotype="Columbia"   |   |
|          | 1..777   | /note="alpha-hydroxynitrile lyase-like protein"                   |   |
|          |  | /codon_start=1  |   |
|          |  | /product="AT5g10300/F18D22_70"                                    |   |
|          |  | /protein_id="AAU10338.1"  |   |
| ORIGIN   |  | /db_xref="GI:20147249"  |   |
|          | /translation="MERKHFVLVYNAYHGMIWYKLPLESAGHRVAVELAASGI<br>DRPFOAVETVEYSKPLETLKSLIPENEEVLVGFSGGINIALAADIIPAKIKVLV<br>FLNAFLPDTTHVPSVLDKYMPPGLGDFEFSHETRNRTMSLKMGPFKMKARLYQ<br>NCPIEDYELAKMLHRQGSFPTEDLSKKKPESEGVSVORVYVMSSEDKAIPCFIRW<br>MIDNFNVSKVYEIDGGDHVMVLSKPKLPFDSLAIATDM" |   |   |
|          | Query Match 85.2%; Score 777; DB 8; Length 777;<br>Best Local Similarity 100.0%; Pred. No. 8.8e-211;<br>Matches 777; Conservative 0; Mismatches 0; Indels 0; Gaps 0;   |   |   |
|          | QY   | 62  | ATGGAGAGAAACATCACTCGTGTAGTTCACACGCTTATCATGGAGCCTGGATCGG 121 |
|          | Db   | 1   | ATGGAGAGAAACATCACTCGTGTAGTTCACACGCTTATCATGGAGCCTGGATCGG 60  |
| QY       | 122  | TACAAGCTCAAGCCCTCTTGAATCAGCGCGGCACCGCTTACTGCTCGAATCGCC 181        |   |
| Db       | 61   | TACAAGCTCAAGCCCTCTTGAATCAGCGCGGCACCGCTTACTGCTCGAATCGCC 120        |   |
| QY       | 182  | GCCTCCGGGATCGACCCACGACCAATCCAGGCGGTGAAACCGTCGACGAATATCCAAA 241    |   |
| Db       | 121  | GCCTCCGGGATCGACCCACGACCAATCCAGGCGGTGAAACCGTCGACGAATATCCAAA 180    |   |
| QY       | 242  | CCGTTGATGAAACCTCAATCTCTTCAGAGAACGAAAGGTAATTCGTTGGATTC 301         |   |
| Db       | 181  | CCGTTGATGAAACCTCAATCTCTTCAGAGAACGAAAGGTAATTCGTTGGATTC 240         |   |
| QY       | 302  | AGCTTCGGAGGATCAACATCGCTCTCGCGCGGACATATTTCCGGCGAAGATTAAGGTT 361    |   |
| Db       | 241  | AGCTTCGGAGGATCAACATCGCTCTCGCGCGGACATATTTCCGGCGAAGATTAAGGTT 300    |   |
| QY       | 362  | CTTGTGTTCTCAACGCTTCTTGCCCGACACAAACCCACGTCCTCTCACGTTCTGGAC 421     |   |
| Db       | 301  | CTTGTGTTCTCAACGCTTCTTGCCCGACACAAACCCACGTCCTTCTCACGTTCTGGAC 360    |   |
| QY       | 422  | AAGTATATGGAGATGCTCGGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAACAAGA 481  |   |
| Db       | 361  | AAGTATATGGAGATGCTCGGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAACAAGA 420  |   |
| QY       | 482  | AATGGGACGATGAGTTTATTTGAAGATGGGACCAAAATTCATGAGGCAAGCTCTTTACCAA 541 |   |
| Db       | 421  | AATGGGACGATGAGTTTATTTGAAGATGGGACCAAAATTCATGAGGCAAGCTCTTTACCAA 480 |   |
| QY       | 542  | AATTGTCCCATAGAGGATTAACGAGCTGGCAAAATTTGTCATAGCGCAAGGTCATTTTTC 601  |   |
| Db       | 481  | AATTGTCCCATAGAGGATTAACGAGCTGGCAAAATTTGTCATAGCGCAAGGTCATTTTTC 540  |   |
| QY       | 602  | ACAGAGGATCTATCAAGAAAGAAAGTTTACGAGGAAGGATATGTTTCGGTGCACAGA 661     |   |
| Db       | 541  | ACAGAGGATCTATCAAGAAAGAAAGTTTACGAGGAAGGATATGTTTCGGTGCACAGA 600     |   |
| QY       | 662  | GTTTACGTAATGAGTAGGAAGCAAGCCATCCCTCGGATTTTCATTCTGTTGATGATT 721     |   |
| Db       | 601  | GTTTACGTAATGAGTAGGAAGCAAGCCATCCCTCGGATTTTCATTCTGTTGATGATT 660     |   |
| QY       | 722  | GATAAATTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTGATGCTC 781 |   |

|  |     |   |     |
|--|-----|---|-----|
| Db   | 661 | GATAATTTCAACGTCCTCGAAAGTCTACGAGATCGATCGGAGATCACATGTTGATGCTC | 720 |
| QY   | 782 | TCCAAACCCCAAAACTCTTTGACTCTCTCTCTGCTATTGCCCACGATTATATGTAA    | 838 |
| Db   | 721 | TCCAAACCCCAAAACTCTTTGACTCTCTCTCTGCTATTGCCCACGATTATATGTAA    | 777 |
| RESULT 5   |     |   |     |
| ATF18D22   |     |   |     |
| LOCUS  |     |   |     |
| Arabidopsis thaliana DNA chromosome 5, BAC clone F18D22 (ESSA project).  |     |   |     |
| ACCESSION  |     |   |     |
| VERSION  |     |   |     |
| KEYWORDS   |     |   |     |
| SOURCE   |     |   |     |
| ORGANISM   |     |   |     |
| Arabidopsis thaliana (thale cress)   |     |   |     |
| Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsi  |     |   |     |
| REFERENCE  |     |   |     |
| AUTHORS  |     |   |     |
| Bevan, M., Zimmermann, W., Gruenewald, A., Wambutt, R., Kalicki, J., Woldmann, P., Smith, A., Bancroft, I., Mewes, H.W., Rudd, S., Lemcke, K. and Mayer, K.F.X.  |     |   |     |
| JOURNAL  |     |   |     |
| REFERENCE  |     |   |     |
| AUTHORS  |     |   |     |
| Bevan, M., Volckaert, G., Grymonprez, B., Voet, M., Robben, J., Bancroft, I., Mewes, H.W., Rudd, S., Lemcke, K. and Mayer, K.F.X.  |     |   |     |
| JOURNAL  |     |   |     |
| REFERENCE  |     |   |     |
| AUTHORS  |     |   |     |
| EU Arabidopsis sequencing project.   |     |   |     |
| Direct Submission  |     |   |     |
| Submitted (05-JUL-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk |     |   |     |
| Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.   |     |   |     |
| FEATURES   |     |   |     |
| source   |     |   |     |
| 1..57180   |     |   |     |
| /organism="Arabidopsis thaliana"   |     |   |     |
| /mol_type="genomic DNA"  |     |   |     |
| /variety="Columbia"  |     |   |     |
| /db_xref="taxon:3702"  |     |   |     |
| /chromosome="5"  |     |   |     |
| misc_feature   |     |   |     |
| 1..20186   |     |   |     |
| /note="Overlap to BAC T31P16, please refer to EMBL:AL356332 for analysis and annotation"   |     |   |     |
| gene   |     |   |     |
| 17729..21213   |     |   |     |
| /gene="F18D22_10"  |     |   |     |
| complement(join(17729..17941,18038..18145,18237..18323,18409..18489,18699..18833,18917..19138,19223..19303,19389..19550,19645..19739,19924..20019,20122..20263,20453..20548,20735..20873,21134..21213))  |     |   |     |
| /gene="F18D22_10"  |     |   |     |
| complement(join(17729..17941,18038..18145,18237..18323,18409..18489,18699..18833,18917..19138,19223..19303,19389..19550,19645..19739,19924..20019,20122..20263,20453..20548,20735..20873,21134..21213))  |     |   |     |
| /gene="F18D22_10"  |     |   |     |
| CDS  |     |   |     |
| complement(join(17729..17941,18038..18145,18237..18323,18409..18489,18699..18833,18917..19138,19223..19303,19389..19550,19645..19739,19924..20019,20122..20263,20453..20548,20735..20873,21134..21213))  |     |   |     |
| /gene="F18D22_10"  |     |   |     |
| note="Contains Glutamine amidotransferases class-II active site AAI-6"   |     |   |     |
| active site AAI-6  |     |   |     |
| 1  |     |   |     |
| /codon_start=1   |     |   |     |
| /product="asparagine synthetase ASN3"  |     |   |     |
| protein id="CAB96680.1"  |     |   |     |
| /db_xref="GI:8953405"  |     |   |     |
| /db_xref="GOA:Q9LFU1"  |     |   |     |
| /db_xref="UniProt/TrEMBL:Q9LFU1"   |     |   |     |
| translation="MCGILAVLGVQVNSQAKRSRIELSRRLHRGPDWMSGLHCYEDC   |     |   |     |
| YLAHERLAIVDPTSGDQPLYNEDKTIATVNGEIYNHKALRENLSKHQFRTGSDCEVI  |     |   |     |

AHLYEHGEEFVDMLDGMFAFVLLDTRDKSFIARDALGITPLYGWGLDGSVNFASE  
MKALSDDCQCFPPGHIYSKQGLRRWYNPMFSEVPSTPYDPLVVRNTEKAV  
IKRLMTDVPFGVLLSGGLDSILVASALRHLEKSEACQWGSKLHTFCIGLKSPDLK  
AGREVADYLGTRHLEHFTVQDIDAIEVIYHVETVDTTIRASTPMFLMSRKIKSL  
GVKMYLSEGSDEIFGGYLFPHKAPNKKEFHEETCRKIKALHOYDCLRANKSTGAWG  
EARVPLDKERTINVAMSIDPEWKMRPDLGRLEKWLGNAPDDEKNPVLPKHILYROK  
EQFSQGVGYSIDGLKHANKHVSETMLNASFVPPDNTPLTKAYYYRTTFEKFPPK  
SAARATVPGGSPVACSTAKAVEWDAWSQNLDPDGRAALGVHVSAYGDEKTEDSRPEK  
LQKJAEKTPAIV"

|        |   |  |
|--------|---|--|
| exon   | complement (17729. .17941)<br>/gene="F18D22_10"<br>/number=1  |  |
| intron | complement (17942. .18037)<br>/gene="F18D22_10"<br>/number=1  |  |
| exon   | complement (18038. .18145)<br>/gene="F18D22_10"<br>/number=2  |  |
| intron | complement (18146. .18236)<br>/gene="F18D22_10"<br>/number=2  |  |
| exon   | complement (18237. .18323)<br>/gene="F18D22_10"<br>/number=3  |  |
| intron | complement (18324. .18408)<br>/gene="F18D22_10"<br>/number=3  |  |
| exon   | complement (18409. .18489)<br>/gene="F18D22_10"<br>/number=4  |  |
| intron | complement (18490. .18698)<br>/gene="F18D22_10"<br>/number=4  |  |
| exon   | complement (18699. .18833)<br>/gene="F18D22_10"<br>/number=5  |  |
| intron | complement (18834. .18916)<br>/gene="F18D22_10"<br>/number=5  |  |
| exon   | complement (18917. .19138)<br>/gene="F18D22_10"<br>/number=6  |  |
| intron | complement (19139. .19222)<br>/gene="F18D22_10"<br>/number=6  |  |
| exon   | complement (19223. .19303)<br>/gene="F18D22_10"<br>/number=7  |  |
| intron | complement (19304. .19388)<br>/gene="F18D22_10"<br>/number=7  |  |
| exon   | complement (19389. .19550)<br>/gene="F18D22_10"<br>/number=8  |  |
| intron | complement (19551. .19644)<br>/gene="F18D22_10"<br>/number=8  |  |
| exon   | complement (19645. .19739)<br>/gene="F18D22_10"<br>/number=9  |  |
| intron | complement (19740. .19923)<br>/gene="F18D22_10"<br>/number=9  |  |
| exon   | complement (19924. .20019)<br>/gene="F18D22_10"<br>/number=10 |  |
| intron | complement (20020. .20121)<br>/gene="F18D22_10"<br>/number=10 |  |
| exon   | complement (20122. .20263)<br>/gene="F18D22_10"<br>/number=11 |  |
| intron | complement (20264. .20452)                                    |  |

|                       |        |                     |       |                                   |
|-----------------------|--------|---------------------|-------|-----------------------------------|
| Query Match           | 68.9%; | Score 628;          | DB 8; | Length 57180;                     |
| Best Local Similarity | 77.6%; | Pred. No. 1.1e-167; |       |                                   |
| Matches               | 912;   | Conservative        | 0;    | Mismatches 0; Indels 264; Gaps 2; |

Qy 1 ATCAGCGTATTCTCAAAACAGCAAGAACCTTAAACCAAAAGAAAAGAAAGCTCTACGGAGACA 60  
|||||  
Db 44418 ATCAGCGTATTCTCAAAACAGCAAGAACCTTAAACCAAAAGAAAAGCTCTACGGAGACA 44477  
|||||

Qy 61 TATGAGAGGAGAAACATCATCTTCGTTAGTTTCACACGCTTATCATGGAGCCTGGATCTG 120

/gene="F18D22\_10"  
/number=11  
complement (20453. .20548)  
/gene="F18D22\_10"  
/number=12  
complement (20549. .20734)  
/gene="F18D22\_10"  
/number=12  
complement (20735. .20873)  
/gene="F18D22\_10"  
/number=13  
complement (20874. .21133)  
/gene="F18D22\_10"  
/number=13  
complement (21134. .21213)  
/gene="F18D22\_10"  
/number=14  
21823. .24163  
/gene="F18D22\_20"  
complement (join(21823. .22209, 22312. .22440, 22508. .22712, 22838. .23739, 23877. .23946, 24033. .24163))  
/gene="F18D22\_20"  
complement (join(21823. .22209, 22312. .22440, 22508. .22712, 22838. .23739, 23877. .23946, 24033. .24163))  
/gene="F18D22\_20"  
/note="strong similarity to non-phototropic hypocotyl 3 - Arabidopsis thaliana, EMBL:AF180390  
Contains Prokaryotic membrane lipoprotein lipid attachment site AA304-314"  
/codon\_start=1  
/product="non-phototropic hypocotyl 3-like protein"  
/protein\_id="CAB96681.1"  
/db\_xref="GI:8953406"  
/db\_xref="GOA:Q9LFU0"  
/db\_xref="UniProt/TreMBL:Q9LFU0"  
/translation="MNSVSAARLKPSFGDAGPVCNKGIIPPSRVVTLANGSEKKDRSW  
YVKSQIPDLSIQVNDITPKAKPEPLISKCYGLSSIELKPSTSENGYHLKLENPFGGA  
DTPEILTKFCYNLPLDNLNVAFLRPLKASLYLWTEFEAGNLIISKTEAFITFVVLAS  
WRDITLVLRSTNLSPWAENLQIVRCCDLLAKACNNNIPEDVDRNERCLYNDIA  
TLDDHFMRVITTKARRAKPQITKIKMYADNLFVINDLEGIKGYLGKNELOF  
SVNRGEESNLGSCQBHETIESLVLPQSGAVSCHFLMLKTSIVYSASPALI  
SDLEKRVMALEDANVCDLLIPNFKNEEQOERVRIPEFFLMHEQQOVLGKPSISKLLD  
NYLAEIAKDPVLPITKFOVLAELMPENAKMCHDGLYRAIDMFLKTHPSLDHRRRLC  
KTNMCEKLSLDACLHAAQNDRPLPRTIVQINTQVLFSEQVKRMMMQDKLPEKEENS  
GGREDKMSRDNEIITKLGELENVKKMSSELQSDYNELOQYERLSSKQKSSHNWGL  
RWQVKKSFQTKREDETRERTRRRRSTGQRTSFRRRMSMS"  
complement (21823. .22209)  
/gene="F18D22\_20"  
/number=1  
complement (22210. .22311)  
/gene="F18D22\_20"  
/number=1  
complement (22312. .22440)  
/gene="F18D22\_20"  
/number=2  
complement (22441. .22507)  
/gene="F18D22\_20"  
/number=2  
complement (22508. .22712)  
/gene="F18D22\_20"  
/number=3  
complement (22713. .22837)  
/gene="F18D22\_20"



|    |       |  |   |       |
|----|-------|--|---|-------|
| Db | 44478 |  | TATGGAGAGAAACATCACTTCGTGTAGTTCAACCGCTTATCATGGAGCCTGGATCTG | 44537 |
| Qy | 121   | GTACAAGCTCAAGCCCTCTCTGAATCAGCCGCGCCACCGGTACTGTGTGAACTCGC     | 180   |       |
| Db | 44538 | GTACAAGCTCAAGCCCTCTCTGAATCAGCCGCGCCACCGGTACTGTGTGAACTCGC     | 44597   |       |
| Qy | 181   | CGCTCTCGGGATCGACCCAGACCAATCCAGGCGCTTGAAACCGTCGACGAATACTCCAA  | 240   |       |
| Db | 44598 | CGCTCTCGGGATCGACCCAGACCAATCCAGGCGCTTGAAACCGTCGACGAATACTCCAA  | 44657   |       |
| Qy | 241   | ACCGTTGATCGAAACCCCTCAATCTCTTCAGAGAACGAAGAGTAATTTCTGGTTGGATT  | 300   |       |
| Db | 44658 | ACCGTTGATCGAAACCCCTCAATCTCTTCAGAGAACGAAGAGTAATTTCTGGTTGGATT  | 44717   |       |
| Qy | 301   | CAGCTTCGGAGGATCAACATCGCTCTCGCCGCGGACATATTTCCGGCGAAGATTAAAGT  | 360   |       |
| Db | 44718 | CAGCTTCGGAGGATCAACATCGCTCTCGCCGCGGACATATTTCCGGCGAAGATTAAAGT  | 44777   |       |
| Qy | 361   | TCCTGTGTTCTCAACGCTCTTTGGCCGACACACCCACGTCCTTCTCACGTTCTGGA     | 420   |       |
| Db | 44778 | TCCTGTGTTCTCAACGCTCTTTGGCCGACACACCCACGTCCTTCTCACGTTCTGGA     | 44837   |       |
| Qy | 421   | C-----   | 421   |       |
| Db | 44838 | CAAGTTAAACAATAAATAACAAAAATAATTTTGATAGTAGTTTATAAATTGAC        | 44897   |       |
| Qy | 422   | -----  | 422   |       |
| Db | 44898 | CAACTTATCTTAATACATGAATAAATAATTTCTGGTTGGTTGATGATATAATAAAGTAT  | 44957   |       |
| Qy | 428   | ATGGAGATGCTCGAGGTTGGGAGATTGTGAGTTTTCATCTCATGAAACAAGAAATGG    | 487   |       |
| Db | 44958 | ATGGAGATGCTCGAGGTTGGGAGATTGTGAGTTTTCATCTCATGAAACAAGAAATGG    | 45017   |       |
| Qy | 488   | ACGATGAGTTTATGGAAGATGGACCAAAATTCATGAAGCGACGCTTTTACCAAAATG    | 547   |       |
| Db | 45018 | ACGATGAGTTTATGGAAGATGGACCAAAATTCATGAAGCGACGCTTTTACCAAAATG    | 45077   |       |
| Qy | 548   | CCCATAG-----   | 554   |       |
| Db | 45078 | CCCATAGAGTATTTGCCACAAATCTCTAAAAATTTGTACGCCAAAAACAATAATCT     | 45137   |       |
| Qy | 555   | -----  | 554   |       |
| Db | 45138 | ATATTCAAACTATGCTGTAAGTTTTCGAATATGAATCATCGTATGAGATGATTATAT    | 45197   |       |
| Qy | 555   | -----  | 554   |       |
| Db | 45198 | TAGTTTTGGTGTGCGTTACATGATGTTACACTGTGAAGGATTACGAGCTGCAAAAT     | 45257   |       |
| Qy | 577   | TTTGCATAGCAGGCTCATTTTTCACAGGATCTATCAAGAAAGAAAGTTTAGCGA       | 636   |       |
| Db | 45258 | TTTGCATAGCAGGCTCATTTTTCACAGGATCTATCAAGAAAGAAAGTTTAGCGA       | 45317   |       |
| Qy | 637   | GGAAGGATATGTTTCGGTGAACAGGTTTACGTAATGAGTAGTGAAGACAAAGCCATCC   | 696   |       |
| Db | 45318 | GGAAGGATATGTTTCGGTGAACAGGTTTACGTAATGAGTAGTGAAGACAAAGCCATCC   | 45377   |       |
| Qy | 697   | CTGCGATTTCAATTCGTTGGATGATGATAATTTCAACGCTTCGAAAGTCTACGAGATCGA | 756   |       |
| Db | 45378 | CTGCGATTTCAATTCGTTGGATGATGATAATTTCAACGCTTCGAAAGTCTACGAGATCGA | 45437   |       |
| Qy | 757   | TGCGGAGATCATCATGGTGTCTCTCCAAACCCGAAACCTTTTGACTCTCTCTGCG      | 816   |       |
| Db | 45438 | TGCGGAGATCATCATGGTGTCTCTCCAAACCCGAAACCTTTTGACTCTCTCTGCG      | 45497   |       |
| Qy | 817   | TATTGCCACGATTTATATGTAATAATCTTTAAGTCCGTTTACTTTTTCTCATCGTTACT  | 876   |       |
| Db | 45498 | TATTGCCACGATTTATATGTAATAATCTTTAAGTCCGTTTACTTTTTCTCATCGTTACT  | 45557   |       |
| Qy | 877   | AATAAAACAAACCCCTTTTTCGGGCAACTTTTCATC                         | 912   |       |

Db

45558

AATAAAACAAACCCCTTTTTCGGGCAACTTTTCATC

45593

RESULT 6

AY202287

LOCUS

DEFINITION

Arabidopsis thaliana sequence flanking Ds3 end of Ds-Trap insertion from line GT6478.

ACCESSION

AY202287

VERSION

AY202287.1

GI:27898241

KEYWORDS

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

1 (bases 1 to 856)

AUTHORS

May,B.P., Simorowski,J., Arroyo,J.-M., Vaughn,M.W., Shen,R., McCombie,W.R. and Martienssen,R.A.

TITLE

Arabidopsis genomic sequences flanking Ds enhancer and gene traps in transgenic lines

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 856)

AUTHORS

May,B.P., Simorowski,J., Arroyo,J.-M., Vaughn,M.W., Shen,R., McCombie,W.R. and Martienssen,R.A.

TITLE

Direct Submission

JOURNAL

Submitted (17-DEC-2002) Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA

COMMENT

http://genetrapp.cehl.org.

FEATURES

Location/Qualifiers

1..856

source

/organism="Arabidopsis thaliana"

/mol\_type="genomic DNA"

/db\_xref="taxon:3702"

/clone="GT6478.DS3.05.08.98.b.888"

/ecotype="Landesberg erecta"

/transgenic

/note="transgenic line GT6478"

1..34

source

/organism="Zea mays"

/mol\_type="genomic DNA"

/db\_xref="taxon:4577"

<1..34

repeat\_region

/transposon="transposon Ds"

ORIGIN

Query Match

37.4%; Score 341.4; DB 12; Length 856;

Best Local Similarity

98.3%; Pred. No. 4e-86;

Matches 353; Conservative

0; Mismatches

5; Indels

1; Gaps

1;

Qy

555

AGGATTACGAGCTGGCAAAATGTTGCATAGGCAAGGTCAATTTTTCACAGAGGATCTAT

614

Db

336

AGGATTACGAGCTGGCAAAATGTTGCATAGGCAAGGTCAATTTTTCACAGAGGATCTAT

395

Qy

615

CAAGAAAGAAAGTTTAGCGAGGAAGGATATGGTTCGGTGCACAGAGTTACGTAATGA

674

Db

396

CAAGAAAGAAAGTTTAGCGAGGAAGGATATGGTTCGGTGCACAGAGTTACGTAATGA

455

Qy

675

GTAGTGAAGACAAAGCCATCCCTCCGATTTTCATTCGTTGGATGATGATTAATTTCAACG

734

Db

456

GTAGTGAAGACAAAGCCATCCCTCCGATTTTCATTCGTTGGATGATGATTAATTTCAACG

515

Qy

735

TCTCCAAAGTCTACGAGATCGATGCGGAGATCACATGTGTATGCTCTCCAAACCCCAA

794

Db

516

TCTCCAAAGTCTACGAGATCGATGCGGAGATCACATGTGTATGCTCTCCAAACCCCAA

575

Qy

795

AACCTTTTGACTCTCTCTGCTATTGGCAGGATTAATGTAATAATCTT-AAAGTCCGT

853

Db

576

AACCTTTTGACTCTCTCTGCTATTGGCAGGATTAATGTAATAATCTTAAAGTCCGT

635

Qy

854

TTTACTTTTTTCTCATCGTTACTATAAATAAACAACCCCTTTTTCGGGGCACTTTTCATC

912

Db

636

TTTACTTTTTTCTCATCGTTACTATAAATAAACAACCCCTTTTTCGGGGCACTTTTCATC

694

```
RESULT 7
AY751530
LOCUS
DEFINITION
Catharanthus roseus protein S mRNA, complete cds.
ACCESSION
AY751530
VERSION
AY751530.1
KEYWORDS
GI:53830669
SOURCE
Catharanthus roseus (Madagascar periwinkle)
ORGANISM
Catharanthus roseus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Gentianales; Apocynaceae; Rauvolfioideae;
Vinceae; Catharanthus.
REFERENCE
1 (bases 1 to 891)
Lemenager,D., Rideau,M. and Clastre,M.
Cloning of the Catharanthus roseus protein S (CrPS) associated with
monoterpenoid indole alkaloid production
Unpublished
JOURNAL
2 (bases 1 to 891)
Lemenager,D., Rideau,M. and Clastre,M.
Direct Submission
AUTHORS
Submitted (15-SEP-2004) Laboratoire de Biologie Molculaire et
Biochimie Vegetale, Faculte de Pharmacie, 31 Avenue Monge, Tours
37200, France
JOURNAL
FEATURES
Location/Qualifiers
1..891
/organism="Catharanthus roseus"
/mol_type="mRNA"
/db_xref="taxon:4058"
48..824
/note="hydrolase; CrPS; alpha/beta fold family protein"
/produit="protein S"
/protein_id="AAU95203.1"
/db_xref="GI:53830670"
/translation="MEVMMKFVTVHGVGHGAWVYVKPRIEAGHRCITAVNLAAGSI
NEKLEVRSSIDYAPLLEVLDSVPENEKVLVGHSGGMTAAVGMKFPNKISLAV
FLNATMPTEENRPSVLBEYTAKTPEAWKQCFSAFGDPPITSLVCCGPETISLTH
LSPIDHALGKTLVRPGLFIEDLKAKEFTBEGFSGVPRVIVIAEDKTIPTPEPQRW
MIENPVKEVKEIKGADHMPFESKFDLSQCLLDIAKHA"
ORIGIN
Query Match 28.0%; Score 255.2; DB 8; Length 891;
Best Local Similarity 56.2%; Pred. No. 1.9e-61;
Matches 479; Conservative 0; Mismatches 373; Indels 0; Gaps 0;
QY 23 AAGAACCCTTACCMAAAGAAAACGCTACGGAGACATATGGAGAGNAACATCACTTC 82
DB 9 ACAGAACTAACATTAGTAGTGAATAGAAAAAATAAGAAATGAAAGTATGAAGCATTTT 68
QY 83 GTGTAGTTTCAACGCTTATCATGGAGCTCGATCTGTCAAGCTCAAGCTCAAGCCCTCTT 142
DB 69 GTGACGGTCCAGCGGTGAGCCAGCGAGCTTGGTCTATTACAGCTGAACCTAGAAAT 128
QY 143 GAATCAGCGCGCCACCGCTTACTGTGTGCAACTCGCGCTCGCGGATCGACCCACGA 202
DB 129 GAGGAGCAGCCACCGCTTGTACGGCGGTAAATTTAGCTGCTTCGTGAATCAACGAAAG 188
QY 203 CCAATCAGCGCGTGAACCGTGCAGGAATCTCMAACCGTTGATCGAAACCCCTCAA 262
DB 189 AAACCTCGAAGAAGTTCGGTCTTCAATTGATTACCGCGCGCGCTTGTGTGAGGTTTGGAT 248
QY 263 TCTCTTCCAGAGAACCAAGAGTAAATCTGTGTGATTTCAGCTCGAGGCATCAACATC 322
DB 249 TCTGTTCTTCAAAATGAGAAGTAAATCTGTGTGAGGCATAGCGCGGTGTATGACGGC 308
QY 323 GCTCTCGCGCCGACATATTTCCGGCGAAGATTAAAGGTTCTTGTGTCTCAACGCTTC 382
DB 309 GCTGTTGGGATGAGAAAATTTCCGAATAAGATCTCTTTAGCTGTCTTTTGAATGCTATA 368
QY 383 TTGCCCCACACRACCACCGTCTCTACGGTCTCTGGACAAGTATATGGAGATGCTTGA 442
```

```
Db 369 ATGCTGATATCCGAAACAGACCACTTATGTTTTGGAGAGTACACTGCAAGACTCCA 428
QY 443 GGTTTGGGAGATTTGTGAGTTTTTCATCTCATGAAACAAGAAATGGGACGATGAGTTTATG 502
Db 429 CCAGAAGCATGGAAGGATTTGCCAGTTCTCAGCCTATGAGATCTCTTATTACATCTTTG 488
QY 503 AAGATGGACCAAAATTCATGAGGACGAGTCTTTTACCAAAATTTGCCCATAGAGATTAC 562
Db 489 GTCTGCGTCTCTGAGTTCAATTTATCCACTCTGTATCACCTCTCCCTTATTGAGGATCAT 548
QY 563 GAGCTGGCAAAATTTGTCATAGGCAAGGTCATTTTTCACAGAGGATCTATCAAGAAA 622
Db 549 GCATTTGGGAAATATTTGTTAAGCCAGGTCATCTATTATAGAGATTTATTGAAGCA 608
QY 623 GAAAAGTTTTAGCGAGGAGATATGGTTTCGGTGCAACGAGTTTACGTAATAGTAGTGAA 682
Db 609 GAAAATTTACAGAGGAGGATTTGGATCAGTACCAAGAGTTTATGTAATAGCTGCTGAA 668
QY 683 GACAAAGCATCTCCCTGCGATTCATTGCTGTTGGATGATGATAATTTCAACGCTCTCGAAA 742
Db 669 GATAAAACAATACCACCTGAAATTCACGTTGGATGATCGAAAAACAACCCAGTTTAAAGGA 728
QY 743 GTCTACGAGATCGATGGGAGATCACATGCTGCTCTCCAAACCCCAAAAACCTCTTT 802
Db 729 GTCAAAGAGATCAAGAGGTGCTGATCATATGCTATGTTCTCTAAACCTGATGAATATCC 788
QY 803 GACTCTCTCTGCTATTGCCACGATTTATATGTAATTAATCTTAAGTCGGTTTACTTTT 862
Db 789 CAATGCTTTTGGACATAGCCCAAGAGCATGCTTTAAGAGTTTACAGACAGAAATTTCTTAC 848
QY 863 TTCTCATCGTTA 874
Db 849 TAAATTTCTTTA 860

RESULT 8
AF269158
LOCUS
DEFINITION
Citrus sinensis ethylene-induced esterase mRNA, complete cds.
ACCESSION
AF269158
VERSION
AF269158.1
KEYWORDS
GI:14279436
SOURCE
Citrus sinensis
ORGANISM
Citrus sinensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Sapindales; Rutaceae; Citrus.
REFERENCE
1 (bases 1 to 1077)
Zhong,G.Y., Goren,R., Riov,J., Sisler,E.C. and Holland,D.
Characterization of an ethylene-induced esterase gene isolated from
Citrus sinensis by competitive hybridization
Physiol. Plantarum 113 (2), 267-274 (2001)
PUBMED
12060305
REFERENCE
2 (bases 1 to 1077)
Zhong,G.Y., Goren,R., Riov,J. and Holland,D.
Direct Submission
AUTHORS
Submitted (18-MAY-2000) Horticulture, Faculty of Agricultural, Food
and Environmental Quality Sciences, The Hebrew University of
Jerusalem, Rehovot 76100, Israel
JOURNAL
FEATURES
Location/Qualifiers
1..1077
/organism="Citrus sinensis"
/mol_type="mRNA"
/db_xref="taxon:2711"
55..858
/note="ELE"
/codon_start=1
/product="ethylene-induced esterase"
/protein_id="AAK58599.1"
/db_xref="GI:14279437"
/translation="WEEVGMEEKHFVLVHGVNHGAWCWLKARLVAGGHRVTAVDL
AASGNMKRIEDVHTFHAYSPLMEVLASLPAEEKVILVGHSLGGVTLLALAGDKPPHK
```

ISVAFVFTAMPDTHRPVLEQYSEKMGKEDDSDWLDTFQSCDASNPISHIMLFG  
EFTIKIYOLCPEDLELAKMLVRPGSMFIDNLKSKFSDGYSKRVYLVCBEDI  
GLPKQFQHWMIQNYPVNEWIEIKGGDHMAWLSDPQKLCDSLQISLKYA"

|        |  |
|--------|--|
| ORIGIN | Query Match<br>Best Local Similarity 27.1%; Score 246.8; DB 8; Length 1077;<br>Matches 456; Conservative 0; Mismatches 312; Indels 12; Gaps 1; |
| QY     | 75 ATCACTTCGTTAGTTCACACGCTTATCATCGAGCCTGGATCTGGTACAGCTCAAGC 134  |
| Db     | 83 AGCATTTGTTCTAGTTTCATGGAGTAACCATGGAGCATGGTGTGGTACAACTGAAAG 142   |
| QY     | 135 CCCTCCTTGAATCAGCCGCGCACCGGTTACTGCTCGAACTCGCGGCTCCGGGATCG 194   |
| Db     | 143 CAAGGCTGGTGGCGGGGGTACCGGGTGAGGCTGTGGACCTAGCGGCTCGGGATCA 202  |
| QY     | 195 ACCCAGACCAATCCAGCGGCTTGAACCGTGCAGCAATCTCCAAACGTTGATCGAAA 254   |
| Db     | 203 ACATGAAGAGAAATTGAGGATGTGCACACATTCATGCATACAGTGAGCCCTTGATGGAGG 262   |
| QY     | 255 CCCTCAAACTCTTCAGAGAACGAGAGGTAATCTGGTTGGATTTCAGCTTCGGAGGCA 314  |
| Db     | 263 TTTTGGGATCAGTTCTCGCGGAGAAAGGTCTACTTGTGCGACACAGCTCTGGGGGG 322   |
| QY     | 315 TCAACATCGCTCTCGCGCGGACATATTTCCGCGGAGGATTAAAGGTTCTTGTGTTCTTCA 374   |
| Db     | 323 TCATTTGGCCCTTGGCGGGACAAATTCACACAAATCTCCGTTGGCTGTTTCGTA 382   |
| QY     | 375 AGCCCTTTTGGCCGACACAAACCCAGTGCCTTCTCACGTTCTGGCAAGTATATGGAGA 434   |
| Db     | 383 CTGCATTTCATGCTGACACACACCGGCCATCTTTGTTTGGAGCAGTATCTCTGAGA 442   |
| QY     | 435 TGCCCTGGA-----GGTTTGGAGATTGTGAGTTTTCATCTCATGAAACAAGAA 482  |
| Db     | 443 AGATGGGAAAGAGGACGACAGCTGTTGGACACTCAATTTTTCACAAATGTGACGCGTCAA 502   |
| QY     | 483 ATGGGACGATGAGTTTATTTGAAGATGGGACCAAAATTCATGAAGGACGCTCTTTACCAA 542   |
| Db     | 503 ATCCATCTCACATTTCCATGCTTTTCGCGCGGAGTTCTTGACTATCAAGATCTATCAGC 562  |
| QY     | 543 ATTGTCCCATAGAGGATACGAGCTGGCAAAATTTGTCATAGGCAAGGTCATTTTCA 602   |
| Db     | 563 TTTGTCTCTGAGGATCTGGAGCTGCCAAGATGTTGCTGAGGCCAGGATCAATGTTTA 622  |
| QY     | 603 CAGAGGATCTATCAAGAAGAAAGTTTATAGCGAGCAAGATATGTTTCGGTGCAACGAG 662   |
| Db     | 623 TAGACAACTTATCGAAGGAAAGTAAGTTTCAGCGATGAAGGATACCGGATCTGTTAAGCGAG 682   |
| QY     | 663 TTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTGCGATTTCATTTCGTTGGATGATG 722   |
| Db     | 683 TTTTATCTTGTATGCGAAGAGGATATTTGGTCTCCCTAAGCAATTTTCAGACTGGATGATCC 742   |
| QY     | 723 ATAATTTCAACGCTCTCGAAAGTCTACAGATCGATGGCGGAGATCAATGGTGATGCTCT 782  |
| Db     | 743 AAAACTATCCGTTAATGAGTGAAGTGAAGATCAAGGGCGGTGATCACATGGCAATGCTTT 802   |
| QY     | 783 CCAAACCCCAAAACTCTTTGACTCTCTCTGCTATTTGCCACCGATATATATATAT 842  |
| Db     | 803 CCGACCCACAGAAACTTTCCGATTTGTCTGTCTCAGATTCTCTTAAGTATGCATAAAATT 862   |

RESULT 9  
AX506625  
LOCUS  
DEFINITION  
Sequence 1320 from Patent WO0216655.  
ACCESSION  
AX506625  
VERSION  
AX506625.1 GI:23387862  
KEYWORDS  
Arabidopsis thaliana (thale cress)  
SOURCE  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

|           |   |
|-----------|---|
| REFERENCE | rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  |
| AUTHORS   | 1 Harper, J.F., Kreps, J., Wang, X. and Zhu, T.   |
| TITLE     | Stress-regulated genes of plants, transgenic plants containing same, and methods of use   |
| JOURNAL   | Patent: WO 0216655-A 1320 28-FEB-2002;<br>The Scripps Research Institute (US); Syngenta Participations AG (CH)  |
| FEATURES  | Location/Qualifiers<br>1..792<br>/organism="Arabidopsis thaliana"<br>/mol_type="unassigned DNA"<br>/db_xref="taxon:3702"  |
| ORIGIN    | Query Match 26.5%; Score 241.6; DB 6; Length 792;<br>Best Local Similarity 59.6%; Pred. No. 1.4e-57;<br>Matches 466; Conservative 0; Mismatches 304; Indels 12; Gaps 3; |
| QY        | 66 AGAGGAAACATCACCTTCGTTAGTTTACAAACGCTTATCATGGAGCCTGGATCTGGTACA 125   |
| Db        | 14 AGAGGAAACATCTTTGTACTAGTACATGGTTCTGTCATGGCGGTGGTCTGGTACA 73   |
| QY        | 126 AGCTCAAGCCCTCCTTGAATCAGCCGCGCACCGGTTACTGCTGTGAACTTCGCGCCT 185   |
| Db        | 74 AGGTTAAGCCGCTGTAGAGCGGTGGGCGCACCGGTAACCTGCTGTGGACTTAGCTGCCT 133  |
| QY        | 186 CCGGATCGACCCA---CGACCAATCCAGCGCGTTGAACCGTCCAGCAATATCTCCAAC 242  |
| Db        | 134 CCGGAATAGACAAACGAGGTGATCACTGACATCCCCACATGCGAAACAATATCGGAGC 193  |
| QY        | 243 CQTTCATCGAAACCCCTCAAACTCTTCCAGAGAACGAGAGGTAATCTTGGTTGGATTCA 302   |
| Db        | 194 CATTGACGAGCTCTTGACCTCATTCGCCAATGATGAAAGGTTGCTCGTTGGTACA 253   |
| QY        | 303 GCTTCGGAGGATCAACATGCTCTCGCGCGCAATATTTCCGCGGAAGATTAAGTTTC 362  |
| Db        | 254 GCTTTGGTGGCTTGAACCTTAGCCATAGCCATGGAAGTTTCCGAAAGAAATCTCTGTG 313  |
| QY        | 363 TTGTGTTCTCAACGCTTCTTGCCGACACAAACCGTGCCTTCACTCGTTCTGGACA 422   |
| Db        | 314 CTGTATTCTTGACTCTTTTCATGCGCGACACGAACTCATCATCTTCGTTCTGGACA 373  |
| QY        | 423 AGTAT-----ATGGAGATGCTCGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAA 476  |
| Db        | 374 AGTTTGGAGCAACATGCTCAAGAGCATGGATGGGACCGCAATTCGAACCTTATGGTT 433   |
| QY        | 477 CAAGAAATGGGACGATGAGTTTATGAAGATGGGACCAAAATTCATGAAGGACGCTCTTT 536   |
| Db        | 434 CAGACAAATTCGGACTGAGTATGTT---TTTAGCCCTGACTTCATGAAGTTGGGTCTCT 490   |
| QY        | 537 ACCAAATTTGCCATAGAGGATACGAGCTGGGCAAAATGTTGATAGGCAAGGTCAT 596   |
| Db        | 491 ACCAGCTTCTCCAGTTGAGGATCTTGAACCTGGGATTACTTTTAAATAGGCCAGGATCG 550   |
| QY        | 597 TTTTTCACAGAGATCTATCAAGAAAGAAAGTTTTCGCGAGGAAAGATATGTTTCGGTGC 656   |
| Db        | 551 TATTTATTAACGATTTATCGAAGATGAAGAACTTCTCGATGAAGGATATGGGTCTGTT 610  |
| QY        | 657 AACGAGTTTACGTAATGAGTAGTGAAGCAAGCCATCCCTTCGCGATTTCAATTCGTTGA 716   |
| Db        | 611 CTCGAGTTTTCATAGTGTGTAAGAGGACAAAGCAATTCAGAAACCGCCAGAGATGGA 670   |
| QY        | 717 TGATTGATTAATTTCAACGCTCTCGAAGTCTAGGATCGATGGCGGAGATCAATCGTTGA 776   |
| Db        | 671 TGATTGATTAATTTTCCGGTGAATTTTAGTGTGGAGATGGAGAGACAGATCATATGCCAA 730  |
| QY        | 777 TGCTCTCAAAACCCCAAAACTCTTTGACTCTCTCTGCTATTGCCACCGATTAATATGT 836  |
| Db        | 731 TGTTCGACGCTCAGCACTCAGTGATTACTTCTGAAAATTCGCGACAAATTCGTTT 790   |
| QY        | 837 AA 838  |



Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGECC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Heuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Becker, J.R. and Theologis, A.

Yamada, K. (SSP/PGECC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGECC) contributed equally to this work as PIs.

Annotation based on July 2002 version of the Arabidopsis genome

# FEATURES

source

Location/Qualifiers

1. .935  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="RAF14-80-F10 (R20356)"  
/ecotype="Columbia"  
/notes="This clone is in a modified pBluescript vector2 (lambda PS) as a BamHI/XhoI insert."

gene

1. .935  
/gene="At2g23620"

5'UTR

1. .43  
/genes="At2g23620"

CDS

44. .835  
/genes="At2g23620"

/codon\_start=1  
/evidence=experimental

/product="putative acetone-cyanohydrin lyase"

/protein\_id="AA022676.1"

/db\_xref="GI:27754457"

/translation="MSERRKQHFVHSGHGWYKVPLEAVGHRVAVDLAA  
SGIDTTRSTDIPTCEQYSEPLKLTSLPNDKVLVGHSEGLNLAIAMEKPKKI  
SVAVFLATMPDTEHSFSLDKFSGNMFQPMGTEPEYGDNSGLSMFSPDFMK  
LGLYQLSPDLELGLMLRPGSLFINDLSKMFNDEGYSVPRFVFCVCKEDKAIFE  
ERQRWIDNFPVNLVMEETDHPMPFCKPQSLDYFLKADKPV"

misc\_difference

341  
/gene="At2g23620"

/note="compared to genomic sequence resulting in an amino

acid sequence difference"

/replaces="g"

836. .935  
/genes="At2g23620"

ORIGIN

Query Match

Best Local Similarity

Matches 467; Conservative

26.5%; Score 241.4; DB 8; Length 935;

Pred. No. 1.7e-57;

Mismatches 306; Indels 12; Gaps 3;

66 AGAGAAACATCACTTCGTTAGTTACACACGCTTATCATGAGCGTGGTACA 125

57 AGAGAAACATCACTTCGTTAGTTACACACGCTTATCATGAGCGTGGTACA 116

126 AGCTCAAGCCCTTCCTTGAATCAGCGCGCACCGCGTTACTGCTGCGAATCTCGCGCCT 185

117 AGGTTAAGCGCTGTAGAGCGGTGGCCACCGGTAATCTGCTGTGAGCTTAGCTGCT 176

186 CCGGGATCGACCA---CGACCAATCCAGCGCGTTGAAACCGTCGACGAATACTCGAAC 242

177 CCGGAATAGACACACGAGTGCATCACTGACATCCCCACATCGCAACAATACTCGGAGC 236

243 CGTTGATCGAACCCTCAATCTCTTCCAGAGACGAGAGTAACTCTGGTTGGATTCA 302

237 CATTTGACGAAGCTCTCTGACCTCATTTGCCAAATGATGAAAAGGTTGTGCTCGTTGGTCA 296

303 GCTTCGGAGGCATCAACATCGCTCTCGCGCGGACATATTTCCGGCGAAGATTAAGGTTTC 362

297 GCTTTGGTGGCTTGAACCTTAGCCATAGCCATGGAAAAGTTTCCCAAAAAAATCTCTGTGC 356  
363 TTGCTTCTCTCAACCCCTTCTTGGCCGACACACCCACGTCGCTTCTCACTGTCGACA 422  
357 CTGTATTTCTTGAATCTCTTATGCGCGGACACCCGACACATCCATCTCTGTCGTCGACA 416  
423 AGTAT-----ATGAGAGATGCTCGAGGTTTGGGAGATTGTGAGATTTTTCATCTCATGAAA 476  
417 AGTTTGGAGACCAACATGCTCAAGAAGCATGATGGCACCAGAAATTCGAACCTTATGGTT 476  
477 CAAGAAATGGGACGATGAGTTTATTTGAAGATGGGACCAAAATTCATGAAGGACGCTCTTT 536  
477 CAGACAAATTCGCGACTGAGTATGTT---TTTTAGCCCTGACTTTCATGAAGTTGGGTCTCT 533  
537 ACCAAAATTTGCTCCCATAGAGGATTACGAGCTGGCAAAAATTTGTCATAGGCAAGGTCAT 596  
534 ACCAGCTTCTCCAGTTGAGGATCTTGAACCTGGATTTACTTTTAATAGGCCAGGATCGT 593  
597 TTTTTCACAGAGGATCTATCAAGAAGAAAAGCTTTTAGCGAGGAGGATATGTTTCGGTGC 656  
594 TATTTATTAAACGATTTATCGAAGATGAAAACCTTCTCGGATGAAGATATGCGTCTGTTTC 653  
657 AACGAGTTTACGTAATGAGTAGTAGAAGACAAAGCCATCCCTCGCGATTTTCATTCGTTGA 716  
654 CTCGAGTTTTCATAGTGTGTAAGAGGACAAAGCAATTCAGAGAAACGCCAGAGATGGA 713  
717 TGATTTGATAATTTCAACGCTCTCGAAAGTCTAGAGATCGATGGCGGAGATCATCATGTTGA 776  
714 TGATTTGATAATTTTCCGGTGAATTTAGTAGGAGATGGAGAGACAGATCATATATATCT 773  
777 TGCTCTCCAAACCCCAAAACCTCTTTGACTCTCTCTGCTATTTGCCACCGGATATATCT 836  
774 TGTCTCGAAGCTCAGCACTCAGTGATTTCTGTAATCTGTAATCTGTAATCTGTAATCTGTA 833  
837 AATAA 841  
834 AATCA 838

RESULT 12

AY485932

LOCUS

DEFINITION

AY485932

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. .1079

/organism="Nicotiana tabacum"

/mol\_type="mRNA"

/cultivar="Xanthi nc [NN]"

/db\_xref="taxon:4097"

42. .824

/note="lipase; SAPP2"

/codon\_start=1

CDS

AY485932  
Nicotiana tabacum salicylic acid-binding protein 2 mRNA linear PLN 07-JAN-2004  
cds.

AY485932  
AY485932.1 GI:40549302

Nicotiana tabacum (common tobacco)

Nicotiana tabacum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

asterids; lamids; Solanales; Solanaceae; Nicotiana.

1 (bases 1 to 1079)

Kumar, D. and Klessig, D.F.

High-affinity salicylic acid-binding protein 2 is required for

plant innate immunity and has salicylic acid-stimulated lipase

activity

Proc. Natl. Acad. Sci. U.S.A. 100 (26), 16101-16106 (2003)

14673096

2 (bases 1 to 1079)

Kumar, D. and Klessig, D.F.

Direct Submission

Submitted (24-NOV-2003) BFI, Tower Rd., Ithaca, NY 14853, USA

Location/Qualifiers

1. .1079

/organism="Nicotiana tabacum"

/mol\_type="mRNA"

/cultivar="Xanthi nc [NN]"

/db\_xref="taxon:4097"

42. .824

/note="lipase; SAPP2"

/codon\_start=1

LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Query Match  
Best Local Similarity  
Matches

QY 33 AACCAAAAGAAAAAGTCTACGAGACATATGAGAGGAAACATCACCTTGGTGTAGTTC 92  
Db 13 AAAAGAAATAACAAGGCATATAAATCAATGAGGAAGGAAACACCTTGTGTAGTAC 72  
QY 93 ACACCGCTTATCATGAGCGCTGGATCTGTACAGCTCAAGCCCTCCTTGAATCAGCG 152  
Db 73 ATGGTGCAATGCCATGGAGGTGGAGTTGGTACAGCTAAAGCCACTGTCTAGAGCTGCAG 132  
QY 153 GCACCGCGCTTACTGCTGCGAACTCGCGCTCGCGGATCGACCAAGCAATCCAGG 212  
Db 133 GCATAAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCACTGATTTGAGAAAATAGAGG 192  
QY 213 CCGTTGAAACCGTCGACGAACTACTCCAAACCGTTGATCGAAACCTCAAACTCTTCCAG 272  
Db 193 AGCTTCGCACACTTTATGATTATCTTTGCCATTGATGAGTTGATGGAATCTCTTCAG 252  
QY 273 AGAAGCAAGAGTAATCTCGTTGGATTCAGCTTCGGAGGATCAACATCGCTCGCG 332  
Db 253 CAGATGAGAAAGTTATATAGTGGGGCATAGTCTTGGTGTATGAATTTGGGACTTGCTA 312  
QY 333 CGACATATTTCCGGCGAAGATTAAGGTTCTGTGTTCTCAACGCTTCTTGCCCGACA 392  
Db 313 TGGAAAGTATCCAAAAGATCTATGCTGCTGTTTCTTGCTGCTTTCATGCTGATT 372  
QY 393 CAACCCAGCTGCTTCTACGTTCTGGCAAGTATATATGAGAGATGCTGGAG-----GTT 446  
Db 373 CTGTTCAACAACCTCTCTTGTGTTGGAACAGTATAATGAGCGGACGCCAGCGAGAATT 432  
QY 447 TGGGAGATTGTCAGTTTTCATCTCATGAAACGAAGAAATGGGACGATGAGTTATTGAAGA 506  
Db 433 GGTGGATACTCAGTTTTTACCATATGGTTCCCTGAAGAGCCACTGACATCCATGTTTT 492  
QY 507 TGGGACCAAAATTCATGAAGGCACGCTTTTACCAAAATTTGCCATAGAGGATTACGAGC 566  
Db 493 TTGGCCCAAGTTCTTGGCTCACAGCTTACAGCTATGCTCTCTGAGGATTTGCAT 552  
QY 567 TGGCAAAATGTTGCATAGGCAAGGTCATTTTTCAGAGAGATCTATCAAGAAGAA 626  
Db 553 TAGCATCATCAATGGTGAGACCAAGCTCTTTGTTTATGGAAGACCTATCGAAGGCCAAGT 612  
QY 627 AGTTAGCAGGAGAGATGTTGCTGTCACAGGATTTACCTAATGAGTAGTGAACA 686  
Db 613 ATTTACAGATGAACGGTTTGGATCAGTGAAGAGAGTTTACATTTGTGCACTGAGGATA 672  
QY 687 AAGCCATCCCTCGCATTTTCAITTCGTTGGATGATTGATAATTTCAACGCTCGAAGTCT 746  
Db 673 AAGGATACCAAGAAATTCAGCGATGGAATTTGACAAATTTGTTGTCATGGAACNA 732  
QY 747 ACAGATCGATCGCGAGATCAATGTTGATGCTCTTCCAAACCCCAAACTCTTTGAT 806  
Db 733 TAGAGATTAAAGTGCTGATCAGTGGCAATGCTATGCGAGGCCCAAACTTTGCGCT 792  
QY 807 CTCTCTCTGCTATTGCGCCGATTTATATGTAAT 839  
Db 793 CTCTCTTGGAAATGTCCTTAATAACACTGAT 825

HEU40402 1078 bp mRNA linear PLN 13-MAR-1996  
Hevea brasiliensis hydroxynitrile lyase (hnl) mRNA, complete cds.  
U40402 GI:1223883  
Hevea brasiliensis (Para rubber tree)  
Hevea brasiliensis  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids 1; Malpighiales; Euphorbiaceae; Crotonoideae;  
Micrandreae; Hevea.  
1 (bases 1 to 1078)  
Hasslacher, M., Schall, M., Hayn, M., Griengl, H., Kohlwein, S.D. and  
Schwab, H. Molecular cloning of the full-length cDNA of (S)-hydroxynitrile  
lyase from Hevea brasiliensis. Functional expression in Escherichia  
coli and Saccharomyces cerevisiae and identification of an active  
site residue  
J. Biol. Chem. 271 (10), 5884-5891 (1996)  
96215058  
8621461  
2 (bases 1 to 1078)  
Hasslacher, M., Schall, M., Hayn, M., Griengl, H., Kohlwein, S.D. and  
Schwab, H.  
Direct Submission  
Submitted (09-NOV-1995) Meinhard Hasslacher, Department of  
Biochemistry, University of Technology Graz, Petersgasse 12/2,  
Graz, 8010, Austria  
Location/Qualifiers  
1. 1078  
/organism="Hevea brasiliensis"  
/mol\_type="mRNA"  
/db\_xref="taxon:3981"  
/tissue\_type="leaves"  
1. 1078  
/gene="hnl"  
44. 817  
/gene="hnl"  
/codon\_start=1  
/product="hydroxynitrile lyase"  
/protein\_id="AAC49184.1"  
/db\_xref="GI:1223884"  
/translation="MAFAHFLVLIHTICHGAWIWHKPLLEALGHKVTALDLAASGVD  
PROIEIGSDPEYSEPLLTLEALPGEKVLVGCESGGLNIAIADKVCYCKIAAAV  
HNSVDPTEHCPISVVDKLMVEVPFDKDTTYFTTKGKEITGLKFTLLRENYTL  
CGPEYELAKMLTRKSLFQNLAKRPFFTKBEGYSIKKIYVTTDQDEIFLPEFQLMQ  
IENYKPKDKVYKVEGGDHKLQLTKTKEIAEILQEVADTYN"  
ORIGIN  
Query Match 24.3%; Score 221.6; DB 8; Length 1078;  
Best Local Similarity 55.3%; Pred. No. 8.2e-52; Indels 0; Gaps 0;  
Matches 431; Conservative 0; Mismatches 349;  
QY 76 TCACCTTCGTGTAGTTTCAACACGCTTATCATGAGCCCTGGATCTGTCACAGCTCAAGCC 135  
Db 55 TCATTTTGTCTTATTCATACCATATGCCAGGTGCATGGATTTGGCACAGCTCAAAACC 114  
QY 136 CCTCTTGAATCAGCCGCCACCGCGTTATCTGCTGTGAACTCGCCGCTCCGGATCGA 195  
Db 115 CCTCTTGGGCACTTGGCCACAAGGTACTGTCACTGGACCTTGACGAAGCGCGTTGA 174  
QY 196 CCCAGCACCATCAGGCCGTTGAAACCGTCGACGAATACTCCAAACCGTTGATCGAATC 255  
Db 175 CCCAAGGCAATGAGGAGATTTGGCTCATTTGATGAGTATTTCTGAACCTTTGTGACGTT 234  
QY 256 CCTCAATCTCTCCAGAGAACGAGGTAATTTCTGTTGGATTTCACTTCAGCTTCGGAGCAT 315  
Db 235 CTTGGAGCACTCCCTCCAGGGGAAAGGTGATTTCTGTTGGCGAGAGCTGTGGAGACT 294  
QY 316 CAACATCGCTCTCGCCGCCGACATATTTCCCGCGGAAGTAAGGTTCTTGTGTTCTCAA 375  
Db 295 CAATATAGCAATTTGCTGCTGATAAATACTGTGAAAAGATTGACGCTGCTGTGTTTCCAAA 354



```
QY 376 CGCTCTTGGCCGACACCAACCCACGTCCTTCTCAGCTTCTGGACAAGTATATGAGAT 435
Db 355 TTCAAGTATTGCCAGACACCGAGCAGTCCCATCTTAGCTGTGATAAAGCTCATGGAGT 414
QY 436 GCCTGGAGGTTTGGGAGATGTGAGTATTTTCATCTCATGAACAAAGAAATGGGACGATGAG 495
Db 415 GTTTCGGACTGGAAAGACACCAAGTATTTTACGTACACTAAAGATGGCAAGGAGATAAC 474
QY 496 TTTTATTGAAGATGGGACCAAAATTCATGAAGGCACGCTCTTTACCAAAATTTGCCATAGA 555
Db 475 TGGATTGAAACTGGGCTTACGCTTCTGAGGGAATTTATATACCTTTGGGCTCTGA 534
QY 556 GGATTACGAGCTGGCAAAATTTGTCATAGCGCAAGGTCATTTTTCACAGAGATCTATC 615
Db 535 GGAATATGAATGGCGAAGATTTGACAAGGAAGGATCATTTTCAAAATATTTTAGC 594
QY 616 AAGAAAGAAAGTTTAGCGGAAGGATATGTTGCGTGGCAACGATTTAGTAATGAG 675
Db 595 TAAGCCAGCAATCTTCACTAAGGAAGGTTACGGATCGATTAAGAAATTTATGTGGAC 654
QY 676 TAGTGAAGACAAAGCCATCCCTCGCATTTTCATTCGTTGGATGATTAATTTCAACGT 735
Db 655 CGACCAAGACGAATATTTTACCTGAATTTCACTCTGCGAAATAGAAAATCTATAAAC 714
QY 736 CTCGAAAGTCTACGAGATCGATGGCGGAGATCATGTTGATGCTCTCCAAACCCCAAA 795
Db 715 AGACAAAGTTTATAAGCTCGAAGTGGAGATCATATAAATTGCAGCTTACAAAGACTAAGGA 774
QY 796 ACTCTTTGACTCTCTCTGCTATTTGCCACCGATTTATATGTAATATCTTAAGTCGGTTT 855
Db 775 GATCGCTGAATTTCTCAAGAGGTGGCTGATACCTATATAATTTGATCTCTTTGAGGCTTTT 834

RESULT 14
A59586
LOCUS A59586
DEFINITION Sequence 1 from Patent WO9703204.
ACCESSION A59586
VERSION A59586.1 GI:3714897
KEYWORDS Hevea brasiliensis (Para rubber tree)
SOURCE Hevea brasiliensis
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Crotonoideae;
Micrandreae; Hevea.
REFERENCE 1
AUTHORS Hasselacher,M., Schall,M., Schwab,H., Hayn,E.M., Kohlwein,S. and
Griengl,H.
TITLE (S)-HYDROXYNITRILLYASE FROM HEVEA BRASILIENSIS
JOURNAL Patent: WO 9703204-A 1 30-JAN-1997;
DSM CHEMIE LINZ GMBH (AT)
FEATURES
source
1..1091
/organism="Hevea brasiliensis"
/mol_type="unassigned DNA"
/db_xref="taxon:3981"
gene 1..1091
/gene="HNL"
mRNA 1..1091
/gene="HNL"
/function="(S)-HYDROXYNITRILLYASE"
ORIGIN
Query Match 24.3%; Score 221.6; DB 6; Length 1091;
Best Local Similarity 55.3%; Pred. No. 8.2e-52;
Matches 431; Conservative 0; Mismatches 349; Indels 0; Gaps 0;
QY 76 TCACCTCGTGTAGTTCAACAGCTTATCATGAGCGCTGATCTGTGTAAGCTCAAGCC 135
Db 68 TCATTTTGTTCATTATCATACCATATATGCCACGGTGCATGATTTGGCAAGCTCAAAAC 127
QY 136 CCTCTCTGAATCAGCGCGCACCGCGTTACTGTGTGAACTCGCGGCTCCGGGATCGA 195
```

```
Db 128 CCTCTCTGAGGCACTTTGGCCACAAGGTTACTGCACTGGACCTTGCAGCAAGCGCGCTTGA 187
QY 196 CCCAGACCAATCCAGGCGGTTGAAACCGTCCAGCAATCTCCAAACCGTTGATCGAAC 255
Db 189 CCCAAGGCCAAATTTGAGGAGATTGGCTCATTTGATGAGTATTTCTGAACCCCTTTGTGACGTT 247
QY 256 CCTCAAAATCTCTTCCAGAGAACGAAGAGGTAAATCTGTTGGATTTCAGCTTCGGAGGCAT 315
Db 248 CTTGAGGCACTCCCTCCAGGGGAAAGGTGATTCTGTTGGGAGAGCTGTGGAGACT 307
QY 316 CAACATGCGCTCTCGCGCGCACATATTTCCGCGCAAGATTAAGGTTCTTGTGTTCTCAA 375
Db 308 CAATATAGCAATTTGCTGTGATAAATACTGTGAAAAGATTTGACGTGCTGTTTTCACAA 367
QY 376 CGCTCTTTCGCCGACACAAACCCACGTCCTCTCAGCTTCCTGGAACAGTATATGAGAT 435
Db 368 TTCAGTATTGCCAGACACCGAGCACTGCCCATCTTACGTCGTGGATAAGCTCATGGAGT 427
QY 436 GCCTGGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAACAAAGAAATGGGACGATGAG 495
Db 428 GTTTCGCCGACTGGAAGACACCAACGTATTTTACGTACACTAAAGATGGCAAGGATTAAC 487
QY 496 TTTATTGAAGATGGGACCAAAATTCATGAAGCAAGTCTTTTACCAAAATTTGCCATAGA 555
Db 488 TGGATTGAAACTGGGCTTACGCTTCTGAGGGAATTTATATACCTTTGGGCTCTGA 547
QY 556 GGATTACGAGCTGCGCAAAATTTGTCATAGGACGAGGTCATTTTTCACAGAGATCTATC 615
Db 548 GGAATATGAACCTGGCGAAGATTTGACAAGGAAGGATCATTTTTCAAAATATTTTAGC 607
QY 616 AAGAAAGAAAGTTTAGCGGAAGGATATGTTTGGTGGCAACAGATTTTACGTAATGAG 675
Db 608 TAAGCGACCATCTTCACTAAGGAGTTACGATCGATTAGAAAATTTATGTGTGGAC 667
QY 676 TAGTGAAGACAAAGCCATCCCTCGCATTTTCATTCGTTGGATGATTAATTTCAACGT 735
Db 668 CGACCAAGACGAATATTTTACCTGAATTTTCACTCTGCGCAATAGAAAATCTATAAAC 727
QY 736 CTCGAAAGTCTACGAGATCGATGGCGGAGATCATGTTGATGCTCTCCAAACCCCAAA 795
Db 728 AGACAAGGTTTATAAGGTCGAAGTGGAGATCATAAATTTGACGCTTACAAAGACTAAGA 787
QY 796 ACTCTTTGACTCTCTCTGCTATTTGCCACCGATTTATATGTAATATCTTAAGTCGGTTT 855
Db 788 GATCGCTGAATTTCTCCAAGAGGTGGCTGATACCTATTAATTCATCTTTGAGGCTTTT 847

RESULT 15
CQ806104
LOCUS CQ806104
DEFINITION Sequence 2515 from Patent WO2004035798.
ACCESSION CQ806104
VERSION CQ806104.1 GI:47111686
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1
AUTHORS Inze,D., de Veylder,L. and Vlieghe,K.
TITLE Identification of novel e2f target genes and use thereof
JOURNAL Patent: WO 2004035798-A 2515 29-APR-2004;
CropDesign N.V. (BE)
FEATURES
Location/Qualifiers
source
1..792
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Query Match 24.0%; Score 218.8; DB 6; Length 792;
```

|  |     |   |     |
|--|-----|---|-----|
| Best Local Similarity 56.6%; Pred. No. 4.8e-51;                |     |   |     |
| Matches 447; Conservative 0; Mismatches 337; Indels 6; Gaps 2; |     |   |     |
| QY   | 55  | GAGACATATGGAGAGAAACATCACTTCGTGTAGTTTCAACACGCTTATCATGGAGCCTG   | 114 |
| Db   | 3   | GAGTGGAGGAGAGAGACACAGTCGTCTCTAGTACATGCTGTGCCATGGCGCCTG        | 62  |
| QY   | 115 | GATCTGGTACAAGCTCAAGCCCTCCTTGAATACGCGGCCACCGGTTACTGTGTGGA      | 174 |
| Db   | 63  | GTGCTGGTACAAGTTAAGCCGAGCTCGAGGCTCTGGCCACCGCTAAACCGCGGTAGA     | 122 |
| QY   | 175 | ACTCGCCGCTCGGGATCGAC---CCACGACCAATCCAGGCCGTGGAACCGTCGACGA     | 231 |
| Db   | 123 | TCTAGCTGCTCCGGTATAGACATGACCCAGGTCAATACAGATATATCCACATGCGAACA   | 182 |
| QY   | 232 | ATACTCCAAACCGTTGATCGAAACCTCAAACTCTCTCCAGAGAACGAGAGGTAAATCT    | 291 |
| Db   | 183 | ATACTCAGAGCCATTGATGCGAGCTAATGACCTCACTACCCAGATGATGAGAAGTTGTGCT | 242 |
| QY   | 292 | GGTTGGATTACGCTTCGGAGGATCAACATCGCTCTCGCCGCCGACATATTTCCGGCGAA   | 351 |
| Db   | 243 | TGTTGGTCACTAGCTTAGGAGGTTTGAGTTTAGCTATGGCCATGGATATGTTCCGACCAA  | 302 |
| QY   | 352 | GATTAAGGTTCTGTGTTCTCTCAACGCCCTCTTGCCCGACACCAACCCAGTCGCTTCTCA  | 411 |
| Db   | 303 | AATCTCTGTTCTGTCTTTGTGACTGCTATGATGCCAGACACCAACACTCACCATCCTT    | 362 |
| QY   | 412 | CGTTCTGGACAAGTATATGGAGATGCTGGAGGTTTGGGAGATTGTGA---GTTTTTCATC  | 468 |
| Db   | 363 | CGTATGGGATAGCTTAAGAAAGAAACCTTCACGAGAGGAATGGTTAGACCCGTGTTTAC   | 422 |
| QY   | 469 | TCAATGAACAAGAAATGGGACGATGATGTTTATGTAAGATGGGACCAAAATTCATGAAGGC | 528 |
| Db   | 423 | GAGCGAGAAACCTGATTTTCTTAGCGAGTTTGGATTTTGGACCAGAAATTCATGCCCAA   | 482 |
| QY   | 529 | ACGTCTTTACCAAAATTTGTCCTATAGAGATTACGAGCTGGCAAAATGTTGCATAGGCA   | 588 |
| Db   | 483 | GAACTTGATACAGTTGTCCTCCAGTCCAGATCTTGAATTTGGCGAAATGTTGGTAGGGC   | 542 |
| QY   | 589 | AGGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTAGCGAGGAAGGATATGG    | 648 |
| Db   | 543 | AAACCATTTGATTAGAAAGATATGCCAGAGAGAGAGCTTCAGTGGAGAGGATACGG      | 602 |
| QY   | 649 | TTCCGTTGCAACGAGTTTACGTAATGAGTAGTGAAAGCAAGCCATCCCTCGCATTTTCA   | 708 |
| Db   | 603 | ATCCGTTACACGTATATTTATTTGATGCGGAAAGGATCTTGTCTCACCCGAAAGATTACCA | 662 |
| QY   | 709 | TCGTTGGATGATTGATAATTTCAACGTCTCGAAAGTCTACGAGATCGATGGCGGAGATCA  | 768 |
| Db   | 663 | GGGATCGATGATCAGCAACTTTCCCCCAAAAGAGTAATGGAGATCAAAAGACGAGATCA   | 722 |
| QY   | 769 | CATGGTGAATGCTCTCCAAACCCCAAAACTCTTTGACTCTCTCTCTGCTATTGCCACCGA  | 828 |
| Db   | 723 | TATGCCAATGTTCTCCAGCCTCAACAACTAATGCTCTCTCTCTTGGAGATTGCAATAA    | 782 |
| QY   | 829 | TTATATGTAA  | 838 |
| Db   | 783 | ATATGCCATA  | 792 |

Search completed: September 12, 2005, 04:45:54  
Job time : 3960.73 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 12, 2005, 21:55:38 ; Search time 609 Seconds  
(without alignments)  
2803.748 Million cell updates/sec

Title: US-10-780-002-2  
Perfect score: 1368  
Sequence: 1 MKEGHFVLVHGAGCHGWSW.....MLCEPQKLCASLLEIAHKYN 260

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 7351250 seqs, 3283620254 residues

Total number of hits satisfying chosen parameters: 14702500

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n model -DEV=xlp  
-Q=/cgn2\_1/USFTO\_spool\_p/US10780002/runat\_11092005\_104749\_26712/app query.fasta\_1.455  
-DB=Published Applications NA -QWTS=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10780002 @CGN 1 1 723 @runat\_11092005\_104749\_26712  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEREQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq.\*  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID                   | Description        |
|------------|-------|-------------|--------|----|----------------------|--------------------|
| 1          | 1368  | 100.0       | 1079   | 21 | US-10-780-002-1      | Sequence 1, Appli  |
| 2          | 837   | 61.2        | 933    | 18 | US-10-424-599-64748  | Sequence 64748, A  |
| 3          | 768   | 56.1        | 792    | 9  | US-09-938-842A-1320  | Sequence 1320, Ap  |
| 4          | 768   | 56.1        | 792    | 11 | US-09-938-842A-1320  | Sequence 1320, Ap  |
| 5          | 768   | 56.1        | 989    | 21 | US-10-780-002-32     | Sequence 32, Appl  |
| 6          | 742   | 54.2        | 1158   | 18 | US-10-424-599-85023  | Sequence 85023, A  |
| 7          | 736   | 53.8        | 1035   | 18 | US-10-425-114-12800  | Sequence 12800, A  |
| 8          | 695   | 50.8        | 792    | 9  | US-09-938-842A-1551  | Sequence 1551, Ap  |
| 9          | 695   | 50.8        | 792    | 11 | US-09-938-842A-1551  | Sequence 1551, Ap  |
| 10         | 695   | 50.8        | 870    | 21 | US-10-780-002-33     | Sequence 33, Appl  |
| 11         | 690   | 50.4        | 792    | 9  | US-09-938-842A-217   | Sequence 217, App  |
| 12         | 690   | 50.4        | 792    | 11 | US-09-938-842A-217   | Sequence 217, App  |
| 13         | 690   | 50.4        | 837    | 21 | US-10-780-002-34     | Sequence 34, Appl  |
| 14         | 680.5 | 49.7        | 792    | 21 | US-10-780-002-35     | Sequence 35, Appl  |
| 15         | 669.5 | 48.9        | 1101   | 19 | US-10-437-963-62048  | Sequence 62048, A  |
| 16         | 665.5 | 48.6        | 819    | 21 | US-10-780-002-37     | Sequence 37, Appl  |
| 17         | 659   | 48.2        | 1069   | 19 | US-10-767-701-9317   | Sequence 9317, Ap  |
| 18         | 654.5 | 47.8        | 834    | 21 | US-10-780-002-39     | Sequence 39, Appl  |
| 19         | 645   | 47.1        | 783    | 21 | US-10-780-002-38     | Sequence 38, Appl  |
| 20         | 640.5 | 46.8        | 771    | 9  | US-09-938-842A-1278  | Sequence 1278, Ap  |
| 21         | 640.5 | 46.8        | 771    | 11 | US-09-938-842A-1278  | Sequence 1278, Ap  |
| 22         | 640.5 | 46.8        | 771    | 21 | US-10-780-002-40     | Sequence 40, Appl  |
| 23         | 635   | 46.4        | 912    | 21 | US-10-780-002-36     | Sequence 36, Appl  |
| 24         | 616   | 45.0        | 653    | 18 | US-10-424-599-52897  | Sequence 52897, A  |
| 25         | 613   | 44.8        | 1007   | 18 | US-10-425-114-21611  | Sequence 21611, A  |
| 26         | 613   | 44.8        | 1018   | 20 | US-10-425-115-119846 | Sequence 119846, A |
| 27         | 600   | 43.9        | 1025   | 18 | US-10-425-114-13774  | Sequence 13774, A  |
| 28         | 600   | 43.9        | 1092   | 20 | US-10-425-115-70617  | Sequence 70617, A  |
| 29         | 590   | 43.1        | 901    | 21 | US-10-780-002-41     | Sequence 41, Appl  |
| 30         | 578   | 42.3        | 996    | 18 | US-10-767-701-15339  | Sequence 15339, A  |
| 31         | 555.5 | 40.6        | 907    | 18 | US-10-425-114-24341  | Sequence 24341, A  |
| 32         | 555.5 | 40.6        | 921    | 20 | US-10-425-115-70618  | Sequence 70618, A  |
| 33         | 550   | 40.2        | 1039   | 18 | US-10-425-114-24225  | Sequence 24225, A  |
| 34         | 550   | 40.2        | 1249   | 20 | US-10-425-115-179879 | Sequence 179879, A |
| 35         | 548   | 40.1        | 1087   | 19 | US-10-437-963-23964  | Sequence 23964, A  |
| 36         | 548   | 40.1        | 1148   | 19 | US-10-437-963-29636  | Sequence 29636, A  |
| 37         | 545   | 39.8        | 786    | 19 | US-10-437-963-62697  | Sequence 62697, A  |
| 38         | 545   | 39.8        | 1173   | 19 | US-10-437-963-89948  | Sequence 89948, A  |
| 39         | 538.5 | 39.4        | 858    | 19 | US-10-437-963-75962  | Sequence 75962, A  |
| 40         | 530.5 | 38.8        | 777    | 21 | US-10-738-922-1      | Sequence 1, Appli  |
| 41         | 530.5 | 38.8        | 777    | 21 | US-10-738-922-2      | Sequence 2, Appli  |
| 42         | 530.5 | 38.8        | 777    | 21 | US-10-738-927-1      | Sequence 1, Appli  |
| 43         | 530.5 | 38.8        | 777    | 21 | US-10-738-927-2      | Sequence 2, Appli  |
| 44         | 527.5 | 38.6        | 920    | 18 | US-10-424-599-128189 | Sequence 128189, A |
| 45         | 526   | 38.5        | 1006   | 20 | US-10-425-115-36145  | Sequence 36145, A  |

ALIGNMENTS

RESULT 1  
US-10-780-002-1  
; Sequence 1, Application US10780002  
; Publication No. US20050034196A1  
; GENERAL INFORMATION:  
; APPLICANT: Klesig, Daniel F.  
; APPLICANT: Kumar, Dharendra  
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN  
; TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, SAMP2, AND METHODS OF USE THEREOF  
; FILE REFERENCE: 3670-PO2652WO  
; CURRENT APPLICATION NUMBER: US/10/780.002  
; CURRENT FILING DATE: 2004-02-17  
; PRIOR APPLICATION NUMBER: PCT/US02/26312  
; PRIOR FILING DATE: 2002-08-16  
; PRIOR APPLICATION NUMBER: 60/312,863  
; PRIOR FILING DATE: 2001-08-16

```
/ NUMBER OF SEQ ID NOS: 58
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 1
/ LENGTH: 1079
/ TYPE: DNA
/ ORGANISM: Nicotiana tabacum
US-10-780-002-1

Alignment Scores:
Pred. No.:      7,15e-164      Length:      1079
Score:          1368.00      Matches:      260
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              21          Gaps:      0

US-10-780-002-2 (1-260) x US-10-780-002-1 (1-1079)

QY 1 MetLysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrp 20
DB 42 ATGAAGGAAGGAAACACCTTTGTTTATACATGCTGTCATGCGATCGAGGTTGGAGTTGG 101
QY 21 TyrLysLeuLysProLeuLeuGluAlaGlyHisLysValThrAlaLeuAspLeuAla 40
DB 102 TACAAGCTAAAGCCACTGCTAGAAGCTGAGGCCATAAGGTTACAGCCCTTGATTTAGCA 161
QY 41 AlaSerGlyThrAspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu 60
DB 162 GTTCTGGCACTGATTTGGAAAAATAGAGGAGCTTCGCACACTTTATGATTATACCTTTG 221
QY 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
DB 222 CCATTGATGAGCTTGATGAATCTCTTTCAGCAGATGAGAGGTTATATATTAGTGGGCAT 281
QY 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
DB 282 AGTCCTGGTGGTATGAATTTGGGACTTGCTATGGAAAAAGTATCCAAAAAGATCATGCT 341
QY 101 AlaValPheLeuAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
DB 342 GCTGTTTCTTGGCTGCTTTCATGCTGATCTGTTTCAACTCTCTCTTGTGTTTGGAA 401
QY 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140
DB 402 CAGTATATAATGAGCGGACGCCAGCCGAGAAATGGTTGGATCTCAGTCTTTTACCATATG 461
QY 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
DB 462 TCCCTTGAAAGCCACTGACATCCATGTTTTTTTGGCCCAAGTTTCTTGCTCACAAAGCTC 521
QY 161 TyrGlnLeuLysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
DB 522 TACCAGCTATGCTCTCTTGGAGATCTTGCTATGATCATCATCTTGGTGAACCAAGCTCT 581
QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
DB 582 TTGTTTATGGAAGACCTATCGAAGGCCAAGATATTTACAGATGAACGGTTTGGATCAGTG 641
QY 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220
DB 642 AAGAGAGTTTACATTGTGTGCACTGAGGATAAAGGCATACCAAGAAATTCACGGCATGG 701
QY 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyValaAspHisMetAla 240
DB 702 CAAATTGACAACATTGGTGTCTCACTGAAGCAATAGAGATTAAAGGTGCTGATCACATGCA 761
QY 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyrAsn 260
DB 762 ATGCTATGCGAGCCCCCAAAACTTTTGGCGCTCTCTCTTGGAAATTTGCCCAATAAACAAC 821

RESULT 2
US-10-424-599-64748
; Sequence 64748, Application US/10424599

; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 64748
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_29480C.1
US-10-424-599-64748

Alignment Scores:
Pred. No.:      3.45e-96      Length:      933
Score:          837.00      Matches:      156
Percent Similarity: 76.86%      Conservative: 40
Best Local Similarity: 61.18%      Mismatches: 55
Query Match:      61.18%      Indels:      4
DB:              18          Gaps:      1

US-10-780-002-2 (1-260) x US-10-424-599-64748 (1-933)

QY 5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyrLysLeuLys 24
DB 43 AAGCACTTTGTTGTTGGTGCATGGGGCATGCCATGGAGCATGGTGTGGTATAAGCTCAAG 102
QY 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAlaSerGlyThr 44
DB 103 CCAGCTTTGGAAATCTGCTGGGCACAGGTCAAGTCTTGACCTTGACCTTCTGGAGGCC 162
QY 45 AspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMetGlu 64
DB 163 AACATGAAGAAATTTGAAGATGTTGACACTTTTTTACAGTATATCTAGAGCTTTGTTGTTT 222
QY 65 LeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGlyGly 84
DB 223 CTATTGGACACAAATTCCTCAATGAAAGAGTAGTCTTAGTGTGTCACAGCTTTGGAGGG 282
QY 85 MetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPheLeu 104
DB 283 CTGACATAGACACTTGCCATGAGGAAATTTCCAGAAAAGGTAGCAGTTGGTGTGTTTCTA 342
QY 105 AlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsnGlu 124
DB 343 ACAGCTTTGCTCCAGAGCTTGAACACACCACCATCTTATGCTTTGGAAAAGTACAGCGAG 402
QY 125 ArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGluGlu 144
DB 403 AGGACCCCGTTAGCTGCTAGTGTGACACTGAATTTGCTCCAGTGGAAAAA----- 456
QY 145 ProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeuCys 164
DB 457 -----ACNACATGTTCTTTGGCCCCCACTTCTTGTCCGACAGCTCTACCACTATCC 510
QY 165 SerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMetGlu 184
DB 511 CCAATTGAGGATTTTGAATTGGCCCAAGACTTTAATAGGCCATCATCACTCTTTCATGGAA 570
QY 185 AspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgValTyr 204
DB 571 GACTTGATGAACAAAAGAACTTCTCCAAAGGGATATGGGTGAGTCCCGCCCTT 630
QY 205 IleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAspAsn 224
DB 631 ATTGTTGCACTGAGGACCTTGCAATTCATTTGGAATATATCAGCTCTTTCATGATCCAAAT 690
```



```
QY 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
Db 193 CCAATTGACGACGCTCGACCTCAATTGCCAAATGATGAAAGGTTGTCTGCTGGTCAC 252
QY 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysValProGlnLysIleTyrAla 100
Db 253 AGCTTTGGTGGCTTGAACATTAGCCATAGCCATGGAAAGTTTCCCGAAAAAATCTCTGTC 312
QY 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
Db 313 GCTGTATTCTTGACTGCTTTCATGCGGACACCGAACACTCACCATCTCTGCTTGGAC 372
QY 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140
Db 373 AAGTTTGGAAAGCAACATGCTCAAGAAGCATGGATGGGCACCGAATTCGAACCTTATGCT 432
QY 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
Db 433 TCAGACAATTCGGACTG---AGTATGTTTTTAGCCCTGACTTCATGAAGTTGGGTCTC 489
QY 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
Db 490 TACCAGCTTCTCCAGTTCAGGATCTTGAACTGGGATTACTTTTAAATGAGGCCAGGATCG 549
QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
Db 550 TTATTATTAAAGCATTTATCGAAGATGAAAAACTTCTCGATCAAGGATATGGTCTGTT 609
QY 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220
Db 610 CTCGAGTGTTCATAGTGTGTAAGAGGACAAAGCAATTCAGAGAAAGCCAGAGATGG 669
QY 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluLysGlyAlaAspHisMetAla 240
Db 670 ATGATTGATAATTTCCGCTGGAATTAGTGATGGAGATGGAGGACAGATCATATGCCA 729
QY 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Db 730 ATGTTCTGCAAGCCTCAGCAACTCAGTGATTACTTCTCTGAAATTCGCGCAAAATTC 786

RESULT 5
US-10-780-002-32
; Sequence 32, Application US/107800002
; Publication No. US20050034196A1
; GENERAL INFORMATION:
; APPLICANT: Klessig, Daniel F.
; APPLICANT: Kumar, Dhirenda
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN
; FILE REFERENCE: 3670-P02652W0
; CURRENT APPLICATION NUMBER: US/10780,002
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: PCT/US02/26312
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/312,863
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 989
; TYPE: DNA
; ORGANISM: Arabidopsis Thaliana
US-10-780-002-32
```

```
Alignment Scores:
Pred. No.: 2,42e-87 Length: 989
Score: 768.00 Matches: 149
Percent Similarity: 71.04% Conservative: 35
Best Local Similarity: 57.53% Mismatches: 73
Query Match: 56.14% Indels: 2
DB: 21 Gaps: 2
```

```
US-10-780-002-2 (1-260) x US-10-780-002-32 (1-989)
QY 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTrpTyr 21
Db 119 AAGAGGAACAACATTTTGTACTAGTACATGGTTCTGTCGCCATGGCGCGTGTGCTGTG 178
QY 22 LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAla 41
Db 179 AAGGTTAAGCGCGTGTAGAGCGGTGGCCACCGCGTAATCTGCTGAGCTTATAGTCC 238
QY 42 SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu 60
Db 239 TCCGGAATAGACACAACGAGGTGCTACTGACATCCCAACATCGCAACATACTCGAG 298
QY 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
Db 299 CCAATTGACGAAGCTCTGACCTCATTTGCCAATGATGAAAGGTTGTCTGCTGGTCAC 358
QY 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
Db 359 AGCTTTGGTGGCTTGAACCTTAGCCATAGCCATGGAAAGTTTCCCGAAAAAATCTCTGTC 418
QY 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
Db 419 GCTGTATTCTTGACTGCTTTCATGCGGACACCGAACACTCACCATCTCTGCTTGGAC 478
QY 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140
Db 479 AAGTTTGGAAAGCAACATGCTCAAGAAGCATGGATGGGCACCGAATTCGAACCTTATGCT 538
QY 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
Db 539 TCAGACAATTCGGACTG---AGTATGTTTTTAGCCCTGACTTCATGAAGTTGGGTCTC 595
QY 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
Db 596 TACCAGCTTCTCCAGTTCAGGATCTTGAACTGGGATTACTTTTAAATGAGGCCAGGATCG 655
QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
Db 656 TTATTATTAAAGCATTTATCGAAGATGAAAAACTTCTCGATCAAGGATATGGGTCTGTT 715
QY 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220
Db 716 CTCGAGTGTTCATAGTGTGTAAGAGGACAAAGCAATTCAGAGAAAGCCAGAGATGG 775
QY 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluLysGlyAlaAspHisMetAla 240
Db 776 ATGATTGATAATTTCCGCTGGAATTAGTGATGGAGATGGAGGACAGATCATATGCCA 835
QY 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Db 836 ATGTTCTGCAAGCCTCAGCAACTCAGTGATTACTTCTCTGAAATTCGCGCAAAATTC 892

RESULT 6
US-10-424-599-85023
; Sequence 85023, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 85023
; LENGTH: 1158
; TYPE: DNA
; ORGANISM: Glycine max
```



```
Db      741 GCTGGATCAATGACGTTCTTAATGATCAAAAGTGCAGATCATGCGAGCTATGTTTAGCAAG 800
      |||:::  :::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      245 ProGlnLysLeuCysAlaSerLeuLeuGluAlaHisLysTyrAsn 260
      |||:::  |||  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Db      801 CCTAGAGACTATTCAATCTCTCCAGAGATAGTACAAATATGAC 848

RESULT 8
US-09-938-842A-1551
; Sequence 1551, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1551
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1551

Alignment Scores:
Pred. No.:      3.6e-78      Length:      792
Score:          695.00      Matches:    138
Percent Similarity: 67.18%      Conservative: 36
Best Local Similarity: 53.28%      Mismatches: 83
Query Match:    50.80%      Indels:     2
DB:              9          Gaps:         2

US-10-780-002-2 (1-260) x US-09-938-842A-1551 (1-792)

Qy      2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyClyTrpSerTrpTyr 21
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Db      13 AAGAGGAAGCAACACTTCGTGTAGTACATGCTGGTGCACGGCGCATGGTGGTAC 72
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      22 LysLeuLysProLeuLeuGluAlaGlyHisLysValThrAlaLeuAspLeuAla 41
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Db      73 AAGGTTAAGCCTCTCTCGAGGCTTTGGGCCATCGGTAAACCGCCTTAGACCTAGCTGCT 132
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      42 SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu 60
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Db      133 TCCGGTATAGACACACACAGGTCATCATCTGACATCTTCTAGATGGAACAATATCTGAG 192
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Db      193 CCATTGATGCAGCTAATGACTTTCATTGCCGAATGATGAGAGGTTGTACTCGTTGGTGCAT 252
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      81 SerLeuGlyClyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Db      253 AGCTTTGGAGGTTTGAGTTTAGCTTTAGCCATGGATAGTTTCCCGATAAATCTCTGTC 312
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Db      313 TCTGCTTCGTGACTGCATTCATGCCCGACACAAACACTCACCCTCGTTCGTCAGGNA 372
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Db      373 AAGTTTGCAAGCAGCATGACACCAAGAGGATGGATGGGCTCTGAGCTCGAGACATATGCT 432
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
```

```
Db      433 TCAGATAATCCGGCTTG---TCTGTGTTCTTCAGACCAGCTTCATGAGCACCGTCTC 489
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Db      490 TACCAACTTTCTCCTGTGGAGGATCTTGAGCTTGGATTGCTTCTAAAGAGGCTAGTTTCA 549
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Db      550 TTGTTTATTAATGAATATATCGAAGATGGAGAACTTTCTCGAGAAAGGGTATGGATCTGTT 609
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGluArgTrp 220
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Db      610 CTCGAGCTTACATGTGTGCAAGAGGACCAACATTATCTCGGAAGACCATCAACGATGG 669
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      221 GlnIleAspAsnIleGlyValThrGluAlaIleGluLysGlyAlaAspHisMetAla 240
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Db      670 ATGATCCATAATTATTCGGCGCAATTTAGTGATTGAGATGGAAGAGACTGATCATATGCCA 729
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Db      730 ATGTTTTCGAAACCTCACTACTAAGTGACCATCTATTGGCAATCGCTGCACAATTTC 786
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::

RESULT 9
US-09-938-842A-1551
; Sequence 1551, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1551
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1551

Alignment Scores:
Pred. No.:      3.6e-78      Length:      792
Score:          695.00      Matches:    138
Percent Similarity: 67.18%      Conservative: 36
Best Local Similarity: 53.28%      Mismatches: 83
Query Match:    50.80%      Indels:     2
DB:              9          Gaps:         2

US-10-780-002-2 (1-260) x US-09-938-842A-1551 (1-792)

Qy      2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyClyTrpSerTrpTyr 21
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Db      13 AAGAGGAAGCAACACTTCGTGTAGTACATGCTGGTGCACGGCGCATGGTGGTAC 72
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      22 LysLeuLysProLeuLeuGluAlaGlyHisLysValThrAlaLeuAspLeuAla 41
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Db      73 AAGGTTAAGCCTCTCTCGAGGCTTTGGGCCATCGGTAAACCGCCTTAGACCTAGCTGCT 132
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      42 SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu 60
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Db      133 TCCGGTATAGACACACACAGGTCATCATCTGACATCTTCTAGATGGAACAATATCTGAG 192
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Db      193 CCATTGATGCAGCTAATGACTTTCATTGCCGAATGATGAGAGGTTGTACTCGTTGGTGCAT 252
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      81 SerLeuGlyClyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Db      253 AGCTTTGGAGGTTTGAGTTTAGCTTTAGCCATGGATAGTTTCCCGATAAATCTCTGTC 312
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Db      313 TCTGCTTCGTGACTGCATTCATGCCCGACACAAACACTCACCCTCGTTCGTCAGGNA 372
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Db      373 AAGTTTGCAAGCAGCATGACACCAAGAGGATGGATGGGCTCTGAGCTCGAGACATATGCT 432
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
```



```

193 CCATTGATGACGTAATACTTCAATTCGCCGAATGATGAGAAGGTTGTACTCGTTGGTCAT 252
QY
81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
253 AGCTTTGGAGGTTTGTAGTTAGCCTTAGCCATGATGAAGTTTCCCGATAAATCTCTGTC 312
QY
101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
313 TCTGTCTCTGAGTGCATTCATTCGCCGACACCAACACTCCACATCGTTGTCGAGGAA 372
QY
121 GlnTyrAsnGluArgThrProAlaGluAsnTyrPheLeuAspThrGlnPheLeuProTyrGly 140
373 AAGTTTTCGAAGCAGCATGACACCAAGGATGGCTCTGAGCTCGAGACATATGTT 432
QY
141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
433 TCAGATAATTCGGCTTG---TCTGTGTTCTTCAGCACCGACTTCATGAAGCACCTCTC 489
QY
161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
490 TACCAACTTCTCTGTGGAGGATCTGAGCTTGGATTGCTTCTAAAGAGGCTTACTTCA 549
QY
181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
550 TTGTTTATTAAATGAATATCGAAGATGGAGAACTTTCTGAGAAAGGATGATGATCTGTT 609
QY
201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTTP 220
610 CCTCGAGCTTACATTTGTGCAAGAGGACCAACATTTCTCGAAGACCATCAACGATGG 669
QY
221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240
670 ATGATCCATAATATTCGGCGAATTTAGTGAATGAGATGGAAGAGACTGATCATATGCCA 729
QY
241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
730 ATGTTTTCGAACCTCAACTACTAAGTGACCATCTATTGGCAATCGCTGACAATTC 786

RESULT 10
US-10-780-002-33
; Sequence 33, Application US/10780002
; Publication No. US20050034196A1
; GENERAL INFORMATION:
; APPLICANT: Klessig, Daniel F.
; APPLICANT: Kumar, Dharendra
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN
; FILE REFERENCE: 3670-PO2652WO
; CURRENT APPLICATION NUMBER: US/10/780,002
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: PCT/US02/26312
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/312,863
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 870
; TYPE: DNA
; ORGANISM: Arabidopsis Thaliana
US-10-780-002-33

Alignment Scores:
Pred. No.: 4.15e-78 Length: 870
Score: 695.00 Matches: 138
Percent Similarity: 67.18% Conservative: 36
Best Local Similarity: 53.28% Mismatches: 83
Query Match: 50.80% Indels: 2
DB: 21 Gaps: 2

US-10-780-002-2 (1-260) x US-10-780-002-33 (1-870)
```

```

QY
2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrTrpSerTyr 21
Db
33 AAGAGGAAGCAACACTTTCGTCTAGTACATGTTGGTGCCACCGCGCATGGTGTGGTAC 92
QY
22 LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAla 41
Db
93 AAGGTTAAGCCCTTCTCGAGGCTTTGGGCCATCGGTAAACCGCCTTAGACCTAGCTGCT 152
QY
42 SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu 60
Db
153 TCGGTATAGACACACACAGGTCATCTACATCTGACAAATATTCGTGAG 212
QY
61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
213 CCATTGATGACGTAATGACTTTCATTCGCGAATGATGAGAAGTTGTACTCGTTGGTCAT 272
QY
81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
Db
273 AGCTTTGGAGGTTTGTAGTTTAGCCTTAGCCATGGATTAAGTTTCCCGATAAATCTCTGTC 332
QY
101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
Db
333 TCTGTCTCTGAGTGCATTCATTCGCCGACACCAACACTCCATCGTTCTGTCGAGGAA 392
QY
121 GlnTyrAsnGluArgThrProAlaGluAsnTyrPheLeuAspThrGlnPheLeuProTyrGly 140
Db
393 AAGTTTTCGAAGCAGCATGACACCAAGGATGGATGGGCTCTGAGCTCGAGACATATGTT 452
QY
141 SerProGluGluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeu 160
Db
453 TCAGATAATTCGGCTTG---TCTGTGTTCTTCAGCACCGACTTCATGAAGCACCGCTCTC 509
QY
161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
Db
510 TACCAACTTCTCTGTGAGGATCTTGAAGCTTGGATTCCTTAAAGAGGCTTACTTCA 569
QY
181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
Db
570 TTGTTTATTAAATGAATATTCGAAGATGGAGAACTTTCTGAGAAAGGATGATGATCTGTT 629
QY
201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTTP 220
Db
630 CCTCGAGCTTACATTTGTGCAAGAGGACCAACATTTCTCGAAGACCATCAACGATGG 689
QY
221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240
Db
690 ATGATCCATAATATTCGGCGAATTTAGTGAATGAGATGGAAGAGACTGATCATATGCCA 749
QY
241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Db
750 ATGTTTTCGAACCTCAACTACTAAGTGACCATCTATTGGCAATCGCTGACAATTC 806

RESULT 11
US-09-938-842A-217
; Sequence 217, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
```

```
; SEQ ID NO 217
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-217

Alignment Scores:
Pred. No.: 1.56e-77 Length: 792
Score: 690.00 Matches: 134
Percent Similarity: 69.92% Conservative: 45
Best Local Similarity: 52.34% Mismatches: 75
Query Match: 50.44% Indels: 2
DB: Gaps: 2

US-10-780-002-2 (1-260) x US-09-938-842A-217 (1-792)

QY 5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTyrTyrLeuLeuLys 24
DB 22 CAACACGTCGTTCTAGTACATGCTGCTGCCATGGCGCTGGTGTGTAACAGGTTAAG 81
QY 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAlaSerGlyThr 44
DB 82 CCGCAGCTCGAGGCTTCTGGCCACCGCTAACCGCGTAGATCTAGCTGCTCCGGTATA 141
QY 45 AspLeu---ArgLysIleGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMet 63
DB 142 GACATGACCCAGTCAATCACAGATATATCCATGCGAACAATACTCAGAGCCATTGATG 201
QY 64 GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 83
DB 202 CAGCTAATGACCTCACTACACGATGATGAGAGGTTGTGTTGTGGTCATAGCTTAGGA 261
QY 84 GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPhe 103
DB 262 GGTTCAGTTAGCTATGCCATGGATATGTTTCCGACCAAAATCTCTGTTCTGCTTT 321
QY 104 LeuAlaAlaPheMetProAspSerValHisAsnSerPheValLeuGluGlnTyrAsn 123
DB 322 GTGACTGTATGATGCCAGACACCAACACTCACCATCCTTCGTATGGATAAGCTAAGA 381
QY 124 GluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeu 163
DB 439 TTTCCTAGCGAGTTTGGATTCTTGACCCAGAATTCATGGCCAAAGAACTTGTATCAGTTG 498
QY 164 CysSerProGluAspLeuAlaSerLeuValArgProSerSerLeuPheMet 183
DB 499 TCTCCAGTCCAAAGATCTTGAATTGGCGAAAATGTTGGTAGGCGCAAACTTGATTAG 558
QY 184 GluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgVal 203
DB 559 AAAGATATGGCAGAGAAGAAGCTTCAGTAGGAAGGATACGGATCCGTTACAGCTATA 618
QY 204 TyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAsp 223
DB 619 TTTATGTATGCGGAAGATCTTGTGTACCCGGAAGATTACCAGCATCGATGATCAGC 678
QY 224 AsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243
DB 679 AACTTTCCCAAAAGAAAGTAATGGAGATCAAAAGACGACAGATCATATGCAATGTTCTCC 738
QY 244 GluProGlnLysCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
DB 739 AAGCCTCAACAACTATGTGCTCTTCTTTGGAGATTGCAATATAATAT 786

RESULT 12
US-09-938-842A-217
; Sequence 217, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
```





```
QY      84 GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPhe 103
      |||:::
Db      259 GGAATACCTGCTCTCGCTCTCAGATATTCCTAGTAAGATTGCTACTATTGTTTTTC 318

QY      104 LeuAlaAlaPheMetProAspSerValHisAsnSerPheValLeuGluGlnTyrAsn 123
      |||:::
Db      319 TTGACAGCTTTTATGCCGACACAGAAACCTACTCTTATGTTTACCAAAAGCTAATC 378

QY      124 GluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGlu 143
      |||:::
Db      379 AGAAGCTTCCACAGAAGGATGTTTGGACACCGTGTGGAACTATGGGAACATGAA 438

QY      144 GluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeu 163
      |||:::
Db      439 TGTCTCTAGAGTTTGCTCTTTTGGACCAAAAGTTTCATGGCCCAAGAAATTTGTATCAACTC 498

QY      164 CysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMet 183
      |||:::
Db      499 TCTCGGTCCAGATCTTGAATGGCAAAATGTTGGTGAGAGTAACCCCATCATTTACA 558

QY      184 GluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgVal 203
      |||:::
Db      559 AATAATCTGGCAGGCAAGAAGCTTTAGTGAGGAAGGTACGGTTACCGTTACACGTATA 618

QY      204 TyrIleValCysThrGluAspLysGlyLysIleProGluGluPheGlnArgTrpGlnIleAsp 223
      |||:::
Db      619 TATATTGTATGTGGAGGACATGGCGGTACCCGAGGATTCACAGTGGTGGATGATCAAG 678

QY      224 AsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243
      |||:::
Db      679 AACTTTCGCCCAAAAGAAATAATGAGATCAAAATGTGCAGATCATATGGCAATGTTCTCC 738

QY      244 GluProGlnLysLeuCysAlaSerLeuGluIleAlaHisLysTyr 259
      |||:::
Db      739 AAGCCTCACAACTATGTCTCTCTCGTGGAGATTGCATGTAATAT 786
```

## RESULT 15

```
US-10-437-963-62048
; Sequence 62048, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 62048
; LENGTH: 1101
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_63420C.1
US-10-437-963-62048
```

```
Alignment Scores:
Pred. No.:      1,07e-74      Length:      1101
Score:          669.50      Matches:      134
Percent Similarity: 67.45%      Conservatives: 38
Best Local Similarity: 52.55%      Mismatches: 82
Query Match:      48.94%      Indels: 1
DB:              19          Gaps: 1
```

US-10-780-002-2 (1-260) x US-10-437-963-62048 (1-1101)

```
QY      5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTrpTyrLysLeuLys 24
      |||:::
Db      81 AAGCACCTCGCTCTGGTCCAGCGCGGTGCATCGCGGCTCGACCTATTTCAAGGTGGCG 140

QY      25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAlaSerGlyThr 44
      |||:::
Db      141 ACGGGCTTCGGTCCGGTCCGGTACCGGTACCGGTACCGGCGCGGACCTCGGCGCTCGGGCGTC 200

QY      45 AspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMetGlu 64
      |||:::
Db      201 GACCGCGCGCGCTGCGCGAGGTGCGGACGTTCGCGACTACCGCGCGCTGCTGTGGCG 260

QY      65 LeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGlyGly 84
      |||:::
Db      261 CTCTCGGCTCCCTCCCTCCAGGGGAGAGGTGCTCTCGTCGGCCACAGGCTCGGGCGGC 320

QY      85 MetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPheLeu 104
      |||:::
Db      321 ATCAAGTTCGGCTGGCGCGGAGCTGTTCCGGGCAAGATCGCGCGCGCTGTGTTCTCTC 380

QY      105 AlaAlaPheMetProAspSerValHisAsnSerPheValLeuGluGlnTyrAsnGlu 124
      |||:::
Db      381 TCGGCTTCATGTCGCGACACACGTCGCGCGCTCGCACGTCTCGAAAGTTTCATCGAG 440

QY      125 ArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGluGlu 144
      |||:::
Db      441 GGGAAAGTGGCGTGGAC---TGGATGGACACGGAGTTTAAGCCGCGAGGATGCGAGGGCAAG 497

QY      145 ProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeuCys 164
      |||:::
Db      498 CTCCCACTTCATGTTGTTTCGGCGCGCAGATCGCCCAAGAAAGGCTGATCGAGCTGTC 557

QY      165 SerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMetGlu 184
      |||:::
Db      558 TCGCCAGAGGACGTCACCTCTCGCGGATCTCTGTAGGGTGAGCTCGATGTTCTGTGGAG 617

QY      185 AspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgValTyr 204
      |||:::
Db      618 GACCTGCAAAAGCAGCAGCCGTTTCAACGAGGGCGGCTACCGGCTCGTCCGGAAGGTGTAC 677

QY      205 IleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAspAsn 224
      |||:::
Db      678 GTGGTCTGTAACACGAGACTCGCCATACCGGAGGGGTTCCAGCGGTGATGATCGGGAAC 737

QY      225 IleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCysGlu 244
      |||:::
Db      738 AGCCCGGTGGATGAGTGAAGGAGATCGATCGCGCGGACACCTCGTGATGCTCTCGAGG 797

QY      245 ProGlnLysLeuCysAlaSerLeuLeuIleAlaHisLysTyr 259
      |||:::
Db      798 CCCGACGAGCTGGCGCGGTGCTCGCCGACATCGCCCGCAAGAGCTAC 842
```

Search completed: September 13, 2005, 00:37:19

Job time : 616 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 12, 2005, 19:33:09 ; Search time 3112 Seconds  
(without alignments)  
3180.177 Million cell updates/sec

Title: US-10-780-002-2

Perfect score: 1368

Sequence: 1 MKEGHFVLVHGACHGWSW.....MLCEPQKLCASLLEIAHKYN 260

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+pn.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US10780002/runat\_11092005\_104747\_26610/app\_query.fasta\_1.455  
-DB=EST -QFMT=fastap -SUFFIX=rest -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=oto -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10780002 @CGN\_1\_1\_5180 @runat\_11092005\_104747\_26610 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

```
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-------------|
| 1          | 1081   | 79.0        | 922    | 7  | CK269872    |
| 2          | 1045.5 | 76.4        | 886    | 7  | CK260266    |
| 3          | 1045.5 | 76.4        | 907    | 7  | CK266902    |
| 4          | 1045.5 | 76.4        | 914    | 7  | CK270870    |
| 5          | 1045.5 | 76.4        | 966    | 7  | CK260265    |
| 6          | 1000.5 | 73.1        | 781    | 4  | BGI26704    |
| 7          | 996.5  | 72.8        | 749    | 5  | BQ514677    |
| 8          | 978.5  | 71.5        | 646    | 2  | AW621893    |
| 9          | 977.5  | 71.5        | 794    | 4  | BM407979    |

|    |       |      |     |   |          |          |           |
|----|-------|------|-----|---|----------|----------|-----------|
| 10 | 965.5 | 70.6 | 784 | 4 | BI434774 | BI434774 | EST537535 |
| 11 | 952.5 | 69.6 | 742 | 7 | CK266903 | CK266903 | EST712981 |
| 12 | 867   | 63.4 | 957 | 7 | CV267710 | CV267710 | WS02032.B |
| 13 | 834   | 61.0 | 639 | 5 | BQ111509 | BQ111509 | EST597085 |
| 14 | 819   | 59.9 | 796 | 7 | CF808374 | CF808374 | P8H8034XM |
| 15 | 798   | 58.3 | 865 | 7 | CN190989 | CN190989 | UCRCS06_0 |
| 16 | 797   | 58.3 | 718 | 5 | BQ853353 | BQ853353 | QGB20812  |
| 17 | 795   | 58.1 | 797 | 6 | CB290370 | CB290370 | UCRCS01_0 |
| 18 | 795   | 58.1 | 821 | 6 | CB290604 | CB290604 | UCRCS01_0 |
| 19 | 792   | 57.9 | 804 | 6 | CB293837 | CB293837 | UCRCS01_0 |
| 20 | 787   | 57.5 | 781 | 7 | CN190896 | CN190896 | UCRCS06_0 |
| 21 | 786   | 57.5 | 712 | 5 | BQ852405 | BQ852405 | QGB17024  |
| 22 | 783   | 57.2 | 811 | 6 | CB292675 | CB292675 | UCRCS01_0 |
| 23 | 780   | 57.0 | 850 | 6 | CB292674 | CB292674 | UCRCS01_0 |
| 24 | 775   | 56.7 | 590 | 4 | BI928709 | BI928709 | EST548598 |
| 25 | 773   | 56.5 | 834 | 7 | CN191164 | CN191164 | UCRCS06_0 |
| 26 | 764   | 55.8 | 747 | 7 | CN189384 | CN189384 | UCRCS06_0 |
| 27 | 762   | 55.7 | 797 | 7 | CF837566 | CF837566 | UCRCS03_0 |
| 28 | 760   | 55.6 | 751 | 7 | CN191146 | CN191146 | UCRCS06_0 |
| 29 | 759   | 55.5 | 535 | 2 | BB433359 | BB433359 | EST399888 |
| 30 | 753   | 55.0 | 835 | 7 | CF830288 | CF830288 | UCRCS01_0 |
| 31 | 750   | 54.8 | 740 | 7 | CN191507 | CN191507 | UCRCS06_0 |
| 32 | 736   | 53.8 | 726 | 7 | CF831499 | CF831499 | UCRCS01_0 |
| 33 | 733   | 53.6 | 733 | 7 | CK110376 | CK110376 | N056C04.P |
| 34 | 728   | 53.2 | 726 | 7 | CF831872 | CF831872 | UCRCS01_0 |
| 35 | 716   | 52.3 | 623 | 5 | BQ854991 | BQ854991 | QGB24M05  |
| 36 | 711   | 52.0 | 828 | 7 | CF838704 | CF838704 | UCRCS03_0 |
| 37 | 708   | 51.8 | 701 | 7 | CN185421 | CN185421 | UCRCS05_0 |
| 38 | 707   | 51.7 | 494 | 2 | B5924354 | B5924354 | EST428123 |
| 39 | 703   | 51.4 | 747 | 1 | AJ805825 | AJ805825 | AJ805825  |
| 40 | 701   | 51.2 | 476 | 1 | AI772823 | AI772823 | EST253923 |
| 41 | 699   | 51.1 | 971 | 5 | BQ481510 | BQ481510 | PV.GBA003 |
| 42 | 697   | 51.0 | 719 | 7 | CN190831 | CN190831 | UCRCS06_0 |
| 43 | 690   | 50.4 | 691 | 7 | CN192311 | CN192311 | UCRCS06_0 |
| 44 | 690   | 50.4 | 911 | 3 | CNS0A981 | BX821370 | Arabidops |
| 45 | 689   | 50.4 | 525 | 1 | AI775631 | AI775631 | EST256731 |

## ALIGNMENTS

CK269872 922 bp mRNA linear EST 03-AUG-2004  
EST715950 potato abiotic stress cDNA library Solanum tuberosum cDNA  
clone POACP55 5' end, mRNA sequence.

CK269872  
CK269872.1 GI:39826850

EST.

Solanum tuberosum (potato)

ORGANISM

VERSION

KEYWORDS

SOURCE

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Other ESTs: EST715951

Unpublished (2003)

Generation of ESTs from abiotic stressed potato tissue

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from the University of Arizona Genomics

Institute via http://genome.arizona.edu/orders/

Seq primer: ART TAG GTG ACA CTA TAG.

Location/Qualifiers

1..922

/organism="Solanum tuberosum"

/mol\_type="mRNA"

/cultivar="Kennebec"

/db\_xref="taxon:4113"

/clone="POACP55"

/tissue\_type="abiotic stress treated leaf and root tissue"

FEATURES  
source

/lab\_host="DH10B-Tona"  
 /clone\_lib="potato abiotic stress cdna library"  
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
 supplier: Solanum tuberosum var. Kennebec plants were  
 grown from cuttings on a 16hr light/8 hr dark cycle at 25  
 C for 3-4 weeks. Abiotic stress conditions were applied to  
 four separate sets of plants. Set 1 involved saturation of  
 the soil with 150 mM NaCl and tissues were harvested at  
 following application of the salt stress (leaves: 2hr,  
 6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).  
 Set 2 were grown under the standard conditions and then  
 were water stressed by withdrawal of further watering  
 applications. Drought stressed plants were harvested after  
 cessation of watering (leaves: 3d, 5d, and 7d; roots:3d  
 and 5d). Set 3 were grown under the standard conditions  
 and then were cold stressed by placement at 4 C. Cold  
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 2d. Set 4 were grown under the standard conditions and  
 then were heat stressed by placement at 35 C. Heat  
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 2d and 4d and heat-stressed roots were harvested at 6 hr,  
 12 hr, 1 d, and 4d. RNA was isolated from all tissues and  
 equal RNA from each tissue and stress was pooled to  
 construct the cdna library. RNA sample."

## ORIGIN

## Alignment Scores:

Pred. No.: 6,07e-117 Length: 922  
 Score: 1081.00 Matches: 207  
 Percent Similarity: 84.50% Conservative: 22  
 Best Local Similarity: 76.38% Mismatches: 30  
 Query Match: 79.02% Indels: 12  
 DB: 7 Gaps: 2

US-10-780-002-2 (1-260) x CK269872 (1-922)

Qy 2 LysGluGlyLysHis-----PheValLeuVal 10  
 Db 4 AAAGAGAAAAAACAAGCCCAAAATGAGGTTATGAGAAACATTTATGTTTGTGTA 63  
 Qy 11 HisGlyAlaCysHisGlyGlyTrpSerTrpValLeuLeuValPheProLeuLeuAla 30  
 Db 64 CATGTGTCATGTCATGGAAGTTGGTGGTATAGCTAAAGCCATTGTTAGAGGCTGCC 123  
 Qy 31 GlyHisLysValThrAlaLeuAspLeuAlaAlaSerGlyThrAspLeuArgLysIleGlu 50  
 Db 124 GACCACACGGTCACCTGCTTGACATGCCGCTCTGGCATTGATTTGAGAAAAATAGAG 183  
 Qy 51 GluLeuArgThrLeuTyraAspTyrThrLeuProLeuMetGluLeuMetGluSerLeuSer 70  
 Db 184 GAACCTCGCACACTTGTGATTATACCGTGCATTCGATTTGAGGATTTATGGAATCCCTTCCA 243  
 Qy 71 AlaAspGluLysValIleLeuValGlyHisSerLeuGlyGlyMetAsnLeuGlyLeuAla 90  
 Db 244 CAAGAAGAGAAAGGTATAGTGGGCATAGTTATGGTGGTATGAAATTTGGACATTGCT 303  
 Qy 91 MetGluLysTyrProGlnLysIleTyrAlaAlaValPheLeuAlaAlaPheMetProAsp 110  
 Db 304 ATGGAAAAATACCAAAAAGATCTTGCTGCTGTTCTTGGCTGCTATTCTGCTGAT 363  
 Qy 111 SerValHisAsnSerPheValLeuGluGlnTyrAsnGluUArgThrProAlaGluAsn 130  
 Db 364 TCTGCTCATATGCTCTCTATGTTTGTATCATGACTTTTGAGAGGACACCAAAAGAGAA 423  
 Qy 131 TrpLeuArgThrGlnPheLeuProTyrGlySerProGluGluProLeuThrSerMetPhe 150  
 Db 424 TGGCTTGACACCAAAATTTGTATCATATGTTTCCCTGAAGAGCCCTCTGACATCCATGTT 483  
 Qy 151 PheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeuCysSerProGluAspLeuAla 170  
 Db 484 TTTGGGCCCCAAGTTCTTGGCTCAACAACCTTTATCATGATTATGCTCTCTCTGAGGATTTGCA 543

Qy 171 LeuAlaSerSerLeuValArgProSerSerLeuPheMetGluAspLeuSerLysAlaLys 190  
 Db 544 TTAGCATCGTCATGGTGAGACCAACCTCTCTGTTTATGAAGATCTATCAAGGTGAAG 603  
 Qy 191 TyrPheThrAspGluArgPheGlySerValLysArgValTyrIleValCysThrGluAsp 210  
 Db 604 TATTTACGACGACGAGGTTATGGATCATGTGAGAAAGTTATATATCTGCATGAGGAT 663  
 Qy 211 LysGlyIleProGluGluPheGlnArgTyrGlnIleAspAsnIleGly----ValThrGlu 229  
 Db 664 AAAGTAACACCAAAAGAAATCCAGCAATGGCAAAATTCACAACAATGGAACCGTCATCGAA 723  
 Qy 230 AlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCysGluProGlnLysLeuCys 249  
 Db 724 GCAAAGGAGATCAAGGTGCTGATCATGTGTAATGCTATGTATGTCGCCCAACAACATTTGT 783  
 Qy 250 AlaSerLeuLeuGluIleAlaHisLysTyrAsn 260  
 Db 784 GACACTCTCTGGAGATTGCCATAAATACAAT 816  
 RESULT 2  
 CK260266  
 LOCUS  
 DEFINITION EST706344 potato abiotic stress cdna library Solanum tuberosum cdna  
 Clone POAB307 5' end, mRNA sequence.  
 ACCESSION CK260266  
 VERSION CK260266.1 GI:39817244  
 KEYWORDS EST.  
 SOURCE Solanum tuberosum (potato)  
 ORGANISM  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum.  
 REFERENCE 1 (bases 1 to 886)  
 AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.  
 TITLE Generation of ESTs from abiotic stressed potato tissue  
 JOURNAL Unpublished (2003)  
 COMMENT Other\_ESTs: EST706342 EST706343 EST706345  
 Contact: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: potato-array@tigr.org  
 Clones can be requested from the University of Arizona Genomics  
 Institute via http://genome.arizona.edu/orders/ .  
 Seg primer: CAG GAA ACA GCT ATG ACC.  
 FEATURES  
 Location/Qualifiers  
 1..886  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec"  
 /db\_xref="taxon:4113"  
 /clones="POAB307"  
 /tissue\_type="abiotic stress treated leaf and root tissue"  
 /lab\_host="DH10B-Tona"  
 /clone\_lib="potato abiotic stress cdna library"  
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
 supplier: Solanum tuberosum var. Kennebec plants were  
 grown from cuttings on a 16hr light/8 hr dark cycle at 25  
 C for 3-4 weeks. Abiotic stress conditions were applied to  
 four separate sets of plants. Set 1 involved saturation of  
 the soil with 150 mM NaCl and tissues were harvested at  
 following application of the salt stress (leaves: 2hr,  
 6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).  
 Set 2 were grown under the standard conditions and then  
 were water stressed by withdrawal of further watering  
 applications. Drought stressed plants were harvested after  
 cessation of watering (leaves: 3d, 5d, and 7d; roots:3d  
 and 5d). Set 3 were grown under the standard conditions  
 and then were cold stressed by placement at 4 C. Cold  
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 2d. Set 4 were grown under the standard conditions and  
 then were heat stressed by placement at 35 C. Heat



```

QY 45 AspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMetGlu 64
Db 162 GATTTGAGAAAAATAGAGAACTTTGCAACACTTGTGTGATTATACCGTGCCATTGATGGAG 221
QY 65 LeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGlyGly 84
Db 222 TTTATGGAATGCTTCCCAAGAGAGAGAGGTCATACCTAGTGGGCGCATGTTATGGTGT 281
QY 85 MetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPheLeu 104
Db 282 ATGAATTTGGCACTTGTCTATGAAAAATACCCAAAAAGATCTTGTGCTGTTACTTG 341
QY 105 AlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsnGlu 124
Db 342 ACTGCTATTCTGCTGATCTACTACATGCTCCTCTATGTTTGGATAAGTACTTGGAG 401
QY 125 ArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGluGlu 144
Db 402 ACGACAACAAAGAGAGATTGGCTTGACACCCCAATTTGTATCATATGGTACCCCTGAAGAG 461
QY 145 ProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeuCys 164
Db 462 CTCGGACATCCATGCTTTTGGGCCCAAGTTCTTGGCTCAAAAACCTTATCATGTTATGC 521
QY 165 SerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMetGlu 184
Db 522 CCTCCTGAGGATCTTTCATTAGCATCGTCTGTTGGTGAGACCAACCTCTCTGTTATCGAA 581
QY 185 AspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgValTyr 204
Db 582 GATCTATCAAGGTGAAGTATTTCACGGACGAAGGGTTTGGATCATATGATGCAAGAAAGTTTAT 641
QY 205 IleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAspAsn 224
Db 642 ATATCTGCACCTGAGGATTAAGTATACCAAAAGAAATTCACGCAATGGCAATTCACAAAC 701
QY 225 IleGly---ValThrGluAlaIleGluLysGlyAlaAspHisMetAlaMetLeuCys 243
Db 702 AATGGAACCGTCATCGAAGCAAGAGATCAAAAGGTACTGATCATATGCAATGCTATGT 761
QY 244 GluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyrAsn 260
Db 762 AAGCCCAACGACTTGTGCGCACTCTCTTGGAGATTGGCCCAATAACATA 812

RESULT 4
CK270870
LOCUS CK270870 914 bp mRNA linear EST 03-AUG-2004
DEFINITION EST1716948 potato abiotic stress cDNA library Solanum tuberosum cDNA
clone POACV79 5' end, mRNA sequence.
ACCESSION CK270870
VERSION CK270870.1 GI:39827848
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE
1. (bases 1 to 914)
Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
Generation of ESTs from abiotic stressed potato tissue
Other ESTs: EST1716949
Unpublished (2003)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES
Location/Qualifiers
1..914
/organism="Solanum tuberosum"

```

```

/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clones="POACV79"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-TonaA"
/clone_lib="potato abiotic stress cDNA library"
/notes="vector: pCMVSPORT6.1; Site_1: EcoRI; Site_2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
12 hr and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 9,59e-113 Length: 914
Score: 1045.50 Matches: 197
Percent Similarity: 85.99% Conservative: 24
Best Local Similarity: 76.65% Mismatches: 35
Query Match: 76.43% Indels: 1
DB: 7 Gaps: 1
US-10-780-002-2 (1-260) x CK270870 (1-914)

```

```

QY 5 LysHisPheValLeuValHisGlyValAcysHisGlyGlyTrpSerTrpTyrLysLeuLys 24
Db 96 AAACACCTTTGTTGGTACATGGTGCAATGCCATGGAAGTTGGTGTGTGTAAGCTAAAG 155
QY 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAlaSerGlyThr 44
Db 156 CCATTTTGTAGAGGTGCGGCCACACGGTCACTGCCCTTGACATGGCGCCCTCTGGCATT 215
QY 45 AspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMetGlu 64
Db 216 GATTTGAGAAAAATAGAGAACTTTGGCACACTTGTGTGATTATACCGTGCCATTGATGGAG 275
QY 65 LeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGlyGly 84
Db 276 TTTATGGAATGCTTCCCAAGAGAGAGAGGTCATACCTAGTGGGCGCATGTTATGGTGT 335
QY 85 MetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPheLeu 104
Db 336 ATGAATTTGGCACTTGTCTATGAAAAATACCCAAAAAGATCTTGTGCTGTTACTTG 395
QY 105 AlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsnGlu 124
Db 396 ACTGCTATTCTGCTGATCTACTACATGCTCCTCTATGTTTGGATAAGTACTTGGAG 455
QY 125 ArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGluGlu 144
Db 456 ACGACAACAAAGAGAGATTGGCTTGACACCCCAATTTGTATCATATGGTACCCCTGAAGAG 515
QY 145 ProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeuCys 164
Db 516 CTTGGACATCATGCTTTTGGGCCCAAGTTCTTGGCTCAAAAACCTTATCATGTTATGC 575

```



Qy 165 SerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMetGlu 184  
 Db 576 CCTCCTGAGGATCTTTTCATTAGCATCGTCTGGTGAGACCAACCTCTCTGTTTATCGAA 635  
 Qy 185 AspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgValTyr 204  
 Db 636 GATCTATCAAGGTGAAGTATTTTCACGACGAAGGGTTTGCATCAGTGAAGAAGTTTAT 695  
 Qy 205 IleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAspAsn 224  
 Db 696 ATATCTGCACTCAGGATAAAGTATCACCAGAAAGATTCAGCAATGGCAATGGAAC 755  
 Qy 225 IleGly---ValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243  
 Db 756 AATGGAACCTCATCGAAGCAAGGAGATCAAGGTACTGATCATATGGCAATGCTATGT 815  
 Qy 244 GluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyrAsn 260  
 Db 816 AAGCCCAACGACTTTGTGCACTCTCTTGGAGATTGCCATAAATACAAT 866

## RESULT 5

CK260265 966 bp mRNA linear EST 03-AUG-2004  
 LOCUS EST706343 potato abiotic stress cdna library Solanum tuberosum CDNA  
 clone POAB307 5' end, mRNA sequence.

ACCESSION CK260265  
 VERSION CK260265.1 GI:39817243  
 KEYWORDS Solanum tuberosum (potato)

Solanum tuberosum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiales; Solanales; Solanaceae; Solanum.

1 (bases 1 to 966)  
 Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.

Generation of ESTs from abiotic stressed potato tissue  
 Unpublished (2003)  
 Other ESTs: EST706342 EST706344 EST706345  
 Contact: Robin Buell

The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org  
 Clones can be requested from the University of Arizona Genomics  
 Institute via <http://genome.arizona.edu/orders/>.

Institute via <http://genome.arizona.edu/orders/>.  
 Seq primer: ATT TAG CTG ACA CTA TAG.

## FEATURES

source

Location/Qualifiers  
 1..966  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec"  
 /db\_xref="taxon:4113"  
 /clone="POAB307"  
 /tissue\_type="abiotic stress treated leaf and root tissue"  
 /lab\_host="DH10B-TonA"  
 /clone\_lib="potato abiotic stress cdna library"  
 /notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
 supplier: Solanum tuberosum var. Kennebec plants were  
 grown from cuttings on a 16hr light/8 hr dark cycle at 25  
 C for 3-4 weeks. Abiotic stress conditions were applied to  
 four separate sets of plants. Set 1 involved saturation of  
 the soil with 150 mM NaCl and tissues were harvested at  
 following application of the salt stress (leaves: 2hr,  
 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).  
 Set 2 were grown under the standard conditions and then  
 were water stressed by withdrawal of further watering  
 applications. Drought stressed plants were harvested after  
 cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
 and 5d). Set 3 were grown under the standard conditions  
 and then were cold stressed by placement at 4 C. Cold  
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 2d. Set 4 were grown under the standard conditions and

then were heat stressed by placement at 35 C. Heat  
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 2d and 4d and heat-stressed roots were harvested at 6 hr,  
 12 hr, 1 d, and 4d. RNA was isolated from all tissues and  
 equal RNA from each tissue and stress was pooled to  
 construct the cDNA library. RNA sample."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1,046-112 Length: 966  
 Score: 1045.50 Matches: 197  
 Percent Similarity: 85.99% Conservatives: 24  
 Best Local Similarity: 76.65% Mismatches: 35  
 Query Match: 76.43% Indels: 1  
 DB: 7 Gaps: 1  
 US-10-780-002-2 (1-260) x CK260265 (1-966)  
 Qy 5 LysHisPheValLeuValHisGlyAlaCysHisGlyTyrSerTrpTyrLysLeuLys 24  
 Db 60 AAACACTTTTGTGATCATGTCATGCCATGGAAGTTGGTGTGATTAAGCTAAAG 119  
 Qy 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaHisSerGlyThr 44  
 Db 120 CCATTGTTAGAGGCTCCCGGCCACACGCTCACTGCCCTTGACATGGCCGCTCTGGCATT 179  
 Qy 45 AspLeuArgLysIleGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMetGlu 64  
 Db 180 GATTTGAGAAAAATAGAGAACTTTGCACACTTGTGATTATACCGTGCCATTGATGGAG 239  
 Qy 65 LeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGlyGly 84  
 Db 240 TTTATGGAATGCTTCCACAGAGAGAGAGGTTCATAGTGGGCGCATGTTATGGTGT 299  
 Qy 85 MetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPheLeu 104  
 Db 300 ATGAATTTGGCACTTGTCTATGGAAAAATACCCAAAAAAGATCTTGTCTGTTACTTG 359  
 Qy 105 AlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsnGlu 124  
 Db 360 ACTGCTATTCGCTGATTTCTACTCAGATGCTCTATGTTTGGATGAAGTACTGGAG 419  
 Qy 125 ArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGluGlu 144  
 Db 420 ACGACACAAAGAGAGAGTGGCTTCACACCAATTTGTTATCATATGTTACCCCTGAAGAG 479  
 Qy 145 ProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeuCys 164  
 Db 480 CCTCGACATCCATGCTTTTGGGCCCAAGTTCTTGGGCTCAAAAACTTTATCATGTTATGC 539  
 Qy 165 SerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMetGlu 184  
 Db 540 CCTCTGAGGATCTTTTCATTAGCATCGTGTGGTGAGACCAACCTCTCTGTTTATCGAA 599  
 Qy 185 AspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgValTyr 204  
 Db 600 GATCTATCAAGGTCAAGTATTTACGACGAGAGGGTTTGGATCAGTGAAGAAGTTTAT 659  
 Qy 205 IleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAspAsn 224  
 Db 660 ATAATCTGCACCTGAGGATAAAGTATCACCAGAAAGAAATTCAGCAATGGCAATGGAAC 719  
 Qy 225 IleGly---ValThrGluAlaIleGluLysGlyAlaAspHisMetAlaMetLeuCys 243  
 Db 720 AATGAACCGTCATCGAAGCAAGGAGATCAAGGTACTGATCATATGGCAATGCTATGT 779  
 Qy 244 GluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyrAsn 260  
 Db 780 AAGCCCAACGACTTTGTGCACTCTCTTGGAGATTGCCATAAATACAAT 830

## RESULT 6

BG126704

LOCUS

BG126704

781 bp

mRNA

linear

EST 31-JAN-2001

DEFINITION EST472350 tomato shoot/meristem Lycopersicon esculentum cDNA clone  
 cTOP13K13 5' sequence, mRNA sequence.  
 ACCESSION BG126704  
 VERSION BG126704.1 GI:12626892  
 KEYWORDS EST.  
 SOURCE Lycopersicon esculentum (tomato)  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 REFERENCE 1 (bases 1 to 781)  
 van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Utterback, T.,  
 Hansen, C., Roming, C. and Tanksley, S.,  
 Generation of ESTs from tomato shoot/meristem tissue  
 Unpublished (2001)  
 CONTACT: CUGI  
 CLEMSON UNIVERSITY  
 CLEMSON UNIVERSITY  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>.  
 Location/Qualifiers  
 1. .781  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cTOP13K13"  
 /tissue\_type="shoot/meristem"  
 /dev\_stage="developing shoots from 4-6wks old plants"  
 /lab\_host="SOLR"  
 /clone\_lib="tomato shoot/meristem"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; Small expanding leaves from the growing tip were  
 taken from greenhouse plants (4-6wks old TA496). Tissue  
 was immediately frozen in liquid nitrogen."  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.62e-107 Length: 781  
 Score: 1000.50 Matches: 188  
 Percent Similarity: 87.1% Conservative: 23  
 Best Local Similarity: 77.6% Mismatches: 30  
 Query Match: 73.1% Indels: 1  
 DB: 4 Gaps: 1  
 US-10-780-002-2 (1-260) x BG126704 (1-781)  
 QY 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTTPSerTrpTyr 21  
 Db 48 AAACAGGAGGACACTTTGTTTGGTACATGTGTGATGCCATGGAGGTGGTGGTGGTAC 107  
 QY 22 LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAla 41  
 Db 108 AAACATAAGCCCTTCTAGAGGTTGCAGGCCACCAAGGTCACAACTCTTGATTGGCAGCC 167  
 QY 42 SerGlyThrAspLeuArgLysLeuGluLeuLeuArgThrLeuTyrAspTyrThrLeuPro 61  
 Db 168 TCTGGTATTGATTGAGAAAAATAGAGCAACTTCACATTTTCATGATTACACTTTGGCA 227  
 QY 62 LeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSer 81  
 Db 228 TTGATGGAATTGATGGAATCTCTTCCACAGAGGAGAAAGTCATCTAGTTGCACATAGT 287  
 QY 82 LeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAla 101  
 Db 288 CTTGGTGGTATGAATTTGGACTTGTATGAAAAATACCCACAAAGATCTATGTGCT 347  
 QY 102 ValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGln 121  
 Db 348 GTTTTCTTGGCTTCATGCTGATTCATCTGCTCTCTCTGCTCTGCTGATCAG 407  
 QY 122 TyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySer 141

Db 408 TATTTTGAGAGGTGCAACCAATGATTTGGTTGGACACCAATTTGTATCATATGTTCC 467  
 QY 142 ProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyr 161  
 Db 468 CATGAAGAGCCTCTGCCATCCATATTTTGGGCCCAAGTCTTGGCTTACAATCTTTAT 527  
 QY 162 GlnLeuCysSerProGluAspLeuAlaLeuSerSerLeuValArgProSerSerLeu 181  
 Db 528 CAATTATGCCCTCTGAGGATTTGCAATGTGATCATCATTTGGGAGAGCAAGTTCITTA 587  
 QY 182 PheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLys 201  
 Db 588 TTCTAGAGATTGTCAAAATCAAGTATTGTGACAGATGAAGGATATGGATCAGTGAAG 647  
 QY 202 ArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGluArgTrpGln 221  
 Db 648 AAAGTTTATATAGTATGCACAGATGATAAATCTCCACAAAAGAGTTTCAAAAATGCCAA 707  
 QY 222 IleAspAsnIle---GlyValThrGluAlaIleGluLysGlyAlaAspHisMetAla 240  
 Db 708 ATTGACCAACATTAAATAGCATCATGAAACAAAAGAAATTTGAAGTCTGTATCATATGGCA 767  
 QY 241 MetLeu 242  
 Db 768 ATGCTA 773  
 RESULT 7  
 BQ514677  
 LOCUS BQ514677  
 DEFINITION EST622092 Generation of a set of potato cDNA clones for microarray  
 analyses mixed potato tissues Solanum tuberosum cDNA clone STMIN36  
 3' end, mRNA sequence.  
 ACCESSION BQ514677  
 VERSION BQ514677.1 GI:21373546  
 KEYWORDS EST.  
 SOURCE Solanum tuberosum (potato)  
 ORGANISM Solanum tuberosum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum.  
 REFERENCE 1 (bases 1 to 749)  
 Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C.,  
 Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and  
 Karamycheva, S.A.  
 Generation of a set of potato cDNA clones for microarray analyses  
 Unpublished (2002)  
 Other ESTs: EST622091  
 Contact: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: [potato-array@tigr.org](mailto:potato-array@tigr.org)  
 This clone can be obtained from the University of Arizona Genomics  
 Institute. Orders can be made through URL:  
<http://genome.arizona.edu/orders/>  
 Seq primer: T7.  
 FEATURES  
 Location/Qualifiers  
 1. .749  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec or Binjite"  
 /db\_xref="taxon:4113"  
 /clone="STMIN36"  
 /tissue\_type="mixed tissues"  
 /lab\_host="SOLR"  
 /clone\_lib="Generation of a set of potato cDNA clones for  
 microarray analyses mixed potato tissues"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; supplier: Combination of untreated and Phytophthora  
 infestans-treated libraries of stolons, leaves, leaflets,  
 axillary buds of stem explants, petioles, germinating  
 eyes, tubers, or roots."  
 ORIGIN







Email: potato-array@tigr.org  
 This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:  
<http://genome.arizona.edu/orders/>  
 Seq primer: M13P-R.

# FEATURES

## Location/Qualifiers

1. .784  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec"  
 /db\_xref="taxon:4113"  
 /clone="PCBU48"  
 /tissue\_type="leaf"  
 /dev\_stage="6 week old"  
 /lab\_host="SOLR"  
 /clone\_lib="P. infestans-challenged potato leaf,  
 compatible reaction"

/notes="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2: XhoI; supplier: Cornell University, Fry lab; sequencing: The Institute for Genomic Research; whole plants were challenged with 20,000 sporangia/ml of the compatible P. infestans isolate US 940480. Leaf tissue was collected at 3, 6, 9, 12, 24, 48, 72 hours after inoculation and frozen in liquid nitrogen immediately upon removal. Kennebec plants showed first symptoms of infection at 48 hours after inoculation. NOTE: We cannot exclude the possibility that this sequence is actually derived from Phytophthora rather than potato."

## ORIGIN

Alignment Scores:  
 Pred. No.: 2,28e-103 Length: 784  
 Score: 965.50 Matches: 188  
 Percent Similarity: 83.73% Conservative: 23  
 Best Local Similarity: 74.60% Mismatches: 40  
 Query Match: 70.58% Indels: 2  
 DB: 4 Gaps: 1

US-10-780-002-2 (1-260) x BI434774 (1-784)

QY 5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTyrTrpLysLeuLys 24  
 DB 30 AAACACTTTGTTGGTACATGCTGTCATGCAAGGAGTTGGTGGTATAGCTAAAG 89  
 QY 25 ProLeuLeuGluAlaGlyHisLysValThrAlaLeuAspLeuAlaSerGlyThr 44  
 DB 90 CCATTGTTAGGCTGCCGCCACACGGTCACCTTGACATGCCCGCCTCTGCATT 149  
 QY 45 AspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMetGlu 64  
 DB 150 GATTGAGAAATAGAGGAACTTTGCACACTTGTGTATATACCGTGCATGTGGAG 209  
 QY 65 LeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 84  
 DB 210 TTTATGGAATGCCCTCCACAGAGGAGAGGTTCATCTAGTGGCGCATGTTATGGT 269  
 QY 85 MetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPheLeu 104  
 DB 270 ATGAATTTGGCATCTGCTATGGAATAATACCCAAAAGATCTTTGTTGCTGTTATG 329  
 QY 105 AlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsnGlu 124  
 DB 330 ACTGCTATTCTGCTGATCTTACTACATGCTCTCTATGTTTGGATAAGTACTTGGAG 389  
 QY 125 ArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGluGlu 144  
 DB 390 ACCGACAAACAAAGAGGATGGCTTGCACCCATTTGTATCATATGTCATCCCTGAAGAG 449  
 QY 145 ProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeuCys 164  
 DB 450 CCTCGGACATCCATGCTTTTGGGCCCAAGTCTTCGGCTCAAAACATTTTATCATGTTATGC 509  
 QY 165 SerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMetGlu 184

DB 510 CCTCTCTAGGATCTTTTCATTAGCATCGTCTGGTGGAGACCAACCTCTCTGTTTATCGAA 569  
 QY 185 AspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgValTyr 204  
 DB 570 GATCTATCAAGGTGAAGTATTTTCACGGACGAGGGTTTGGATCAGTGAAGAAGTTTAT 629  
 QY 205 IleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAspAsn 224  
 DB 630 ATATCTGCACCTGAGGATTAAGTATCACCAAAAGATTTCCAGCATGGCAATTGACAAC 689  
 QY 225 IleGly---ValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243  
 DB 690 AATGGAACCGTCTCATCGAACAAAGGAGATCAAGGTACTGATCATATGCGCAATGCTATGT 749  
 QY 244 GluProGlnLysLeuCysAlaSerLeuLeuGluIle 255  
 DB 750 AAGCCCAA-CGACTTTGTGCGCACTCTCTTGGAGATT 784

## RESULT 11

CK266903/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Solanum tuberosum (potato)

Solanum tuberosum

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CK266903 742 bp mRNA linear EST 03-AUG-2004  
 EST712981 potato abiotic stress cDNA library Solanum tuberosum cDNA  
 clone POAC751 3' end, mRNA sequence.  
 CK266903  
 CK266903.1 GI:39823881

EST.

Solanum tuberosum (potato)

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiids; Solanales; Solanaceae; Solanum.

Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.

Generation of ESTs from abiotic stressed potato tissue

Unpublished (2003)

Other ESTs: EST712980

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from the University of Arizona Genomics

Institute via <http://genome.arizona.edu/orders/>

Seq primer: GPA ATA CGA CTC ATA GGG C.

## FEATURES

### source

1. .742  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec"  
 /db\_xref="taxon:4113"  
 /clone="POAC751"  
 /tissue\_type="abiotic stress treated leaf and root tissue"  
 /lab\_host="DH10B-Tona"  
 /clone\_lib="potato abiotic stress cDNA library"  
 /note="Vector: pCMVSPORT6.1; Site.1: EcoRI; Site.2: NotI;  
 supplier: Solanum tuberosum var. Kennebec plants were  
 grown from cuttings on a 16hr light/8 hr dark cycle at 25  
 C for 3-4 weeks. Abiotic stress conditions were applied to  
 four separate sets of plants. Set 1 involved saturation of  
 the soil with 150 mM NaCl and tissues were harvested at  
 following application of the salt stress (leaves: 2hr,  
 6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).  
 Set 2 were grown under the standard conditions and then  
 were water stressed by withdrawal of further watering  
 applications. Drought stressed plants were harvested after  
 cessation of watering (leaves: 3d, 5d, and 7d; roots:3d  
 and 5d). Set 3 were grown under the standard conditions  
 and then were cold stressed by placement at 4 C. Cold  
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
 2d. Set 4 were grown under the standard conditions and  
 then were heat stressed by placement at 35 C. Heat  
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,



```
Db 159 AAGAAATCAGAAGCATTTTGTCTAGTACATGGAGCTTGCATGGCGCTTGGTCTGGCAA 218
Qy 22 LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAla 41
Db 219 AAGTTCAAAACGCTGCTGTAGTACAGCAAGTAACAGGCTCACGGTGTCTGACCTTGTCTGT 278
Qy 42 SerGlyThrAspLeuArgLysLysLeuGluLeuArgThrLeuLysAspTyrThrLeuPro 61
Db 279 TCAGGGCCCAACATGAAGCAATCCAAGATGTAGAACAACACTTGATGAATATACGGAGCCT 338
Qy 62 LeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValLeuValGlyHisSer 81
Db 339 TTGTTAGAGTTTCTGGCTCATTAACAACCAAGAGAGAGTCACTTCTAGTAGGCGCACAGC 398
Qy 82 LeuGlyCysMetAsnLeuLeuAlaMetGluLysTyrProGlnLysLysLeuAlaAla 101
Db 399 CTAGGAGGTTTGTAGTTGGCTCTGTATGGAAGTTCCAGAGAAAGATGCTGTGCT 458
Qy 102 ValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGln 121
Db 459 GTTTTCTTATCAGCTTTCATGCCAGATACACACAAAGCCATCATTTGTCTTGGATCAG 518
Qy 122 TyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySer 141
Db 519 TATAACGAGAGGACCCCGCGGATTCCTGTTGGACACTCAATTTTACCATACAGCACT 578
Qy 142 ProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyr 161
Db 579 TCTCAAAAGTCATCTCAACAATGTCTTTTGGACCCAAATTTCTATCTCCCAAGCTCTAT 638
Qy 162 GlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerLeu 181
Db 639 CAGCTAAGCCACCTCAGGATCTTTGAGCAAGCAAGACTATGTATAGGCGAGATCACATG 698
Qy 182 PheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLys 201
Db 699 TTTCTGTATGATTTGTCAAAGGCAACAGTTTCTCCACGACGGCTATGGTCAGTCAAA 758
Qy 202 ArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGln 221
Db 759 CGAGTGTATGTTATCTCGCATGAAGATTTAGCGATACCCAGAGAGTTTCAAGCGCTGGATG 818
Qy 222 IleAspAsnIleGlyValThrGluAlaIleGluLysGlyAlaAspHisMetAlaMet 241
Db 819 ATTGAANAACGCTCTGTGAGAAAGTTATGGAATTTGAAGGTGCAGACCATATGGTTATG 878
Qy 242 LeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Db 879 TTCAGCAAGCCACAGAAGACTCTTCCATTGTCTCTCAGAGATAGCAAAATAACAT 932

RESULT 13
BQ111509
LOCUS
DEFINITION
5' end, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Solanum tuberosum (potato)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 639)
REFERENCE
AUTHORS
Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C.,
Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and
Karamycheva, S.A.
Generation of a set of potato cDNA clones for microarray analyses
Unpublished (2002)
On Apr 17, 2002 this sequence version replaced gi:20163471.
Contact: Robin Buell
The Institute for Genomic Research
```

```
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T3.
FEATURES
Location/Qualifiers
source
1..639
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec or Binjete"
/db_xref="taxon:4113"
/clone="STMCB21"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/clone_lib="mixed potato tissues"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Combination of untreated and Phytophthora
infestans-treated libraries of stolons, leaves, leaflets,
axillary buds of stem explants, petioles, germinating
eyes, tubers, or roots."
```

## ORIGIN

```
Alignment Scores:
Pred. No.: 6,32e-88 Length: 639
Score: 834.00 Matches: 157
Percent Similarity: 86.57% Conservative: 17
Best Local Similarity: 78.11% Mismatches: 27
Query Match: 60.96% Indels: 0
DB: 5 Gaps: 0
```

US-10-780-002-2 (1-260) x BQ111509 (1-639)

```
Qy 5 LysHisPheValLeuValHisGlyAlaCysHisGlyTyrSerTyrLysLeuLys 24
Db 35 AAACACTTTTGTGGTACATGGTGCATGCCATGGAAGTTGGTGTATAGTAAAG 94
Qy 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAlaSerGlyThr 44
Db 95 CCAATGTTAGAGCTGCGGCCACACGGTCACTGCCCTTGACATGCCGCTCTGGCATT 154
Qy 45 AspLeuArgLysIleGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMetGlu 64
Db 155 GATTTGAGAAAATAGAGAACTTTTGGCACACTTGTGATTATACCGTGCATTGATGGAG 214
Qy 65 LeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 84
Db 215 TTTATGGAATGCCTTCCAAAGAGAGAGGTCATACCTAGTGGGCAATGTTATGTTGCT 274
Qy 85 MetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPheLeu 104
Db 275 ATGAATTTGGCACTTGTCTATGGAANAATACCCAAAAGAGATCTTGTGCTGTTACTTG 334
Qy 105 AlaAlaPheMetProAspSerValHisAsnSerPheValLeuGluGlnTyrAsnGlu 124
Db 335 ACTGCTATTCTGCCTGATTCTACTCACATGCTCTCTATGTTTGGATAAGTACTTGGAG 394
Qy 125 ArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGluGlu 144
Db 395 ACGACAACAAGAGAGGATTGGCTTGACACCCCAATTTGTATCATATGGTACCCCTGAAGAG 454
Qy 145 ProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeuCys 164
Db 455 CCTCGGACATCCATCTTTTGGGCCCAAGTTCTTGGCTCAAAAACCTTTATCAGTTATGC 514
Qy 165 SerProGluAspLeuAlaAlaSerSerLeuValArgProSerSerLeuPheMetGlu 184
Db 515 CCTCTCTAGGATCTTTCAATAGCATCGTGTGTTGGTGAGACCAACCTCTCTGTTATCGAA 574
Qy 185 AspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgValTyr 204
Db 575 GATCTATCAAGGTTGAAGTATTTTACCGGACGAAAGGTTTGGATCAGTGAAGAAGTTTAT 634
```



```

/tissue_type="Rind"
/dev_stage="Commercially producing trees"
/lab_host="E. coli TJC121"
/clone_lib="Washington Navel Orange Stored Fruit Rind cDNA Library"

/note="vector: Lambda Uni-ZAP XR, excised phagemid; Site 1: EcoRI; Site 2: XhoI; Fruits were collected January-March 2003 (Federick, Roose lab; Focht, Sleavert & Robinson, Arpaia lab). Four samples related to storage conditions were produced: 1) fresh-picked in Mantone (Arnott Brothers Enterprises, Mantone, CA), 2) after 21 days storage at 5C at Kearney then transported to UC Riverside on ice, 3) after 5 additional days storage at 11C at Kearney, sampled immediately, 4) fruit grown in southern CA were obtained from Redlands Foothill Packing House after commercial packing, X-ray irradiated at 300 Gy by Surebeam, then stored 1 day at ambient temperature. Rind tissue (juice vesicles) were collected. Tissues were snap frozen and then stored at -80C until further processing. Fenton (Close lab) purified RNA by the phenol method described in J. Japanese Soc. Hort. Sci. 1996. 64 (4): 809-814, purified poly(A) mRNA using a PolyAtrack mRNA Isolation System IV (Promega), produced a primary cDNA library using a lambda ZAP XR cDNA Synthesis Kit (Stratagene), then mass-excised one million pfu from the primary library to produce a phagemid population. Phagemids were plated, plasmid DNA purified. cDNA clones archived, and DNA sequences determined bi-directionally using an ABI3730 at the Arizona Genomics Institute, University of Arizona (Kim, Kudrna, Stum, Yost, Wing). Chromatogram files were downloaded by FTP to UC Riverside (by Close), then processed at UC Riverside (by Wanamaker, Close lab) using the HARVEST pipeline (http://harvest.ucr.edu) to remove vector and cloning oligo sequences and various contaminants, and to trim to a high quality region. Sequences that retained a phred 17 region of at least 100 bases were deposited to GenBank."

```

## ORIGIN

|                         |          |                  |
|-------------------------|----------|------------------|
| Alignment Scores:       |          |                  |
| Pred. No.:              | 1,828-83 | Length:          |
| Score:                  | 798.00   | Matches:         |
| Percent Similarity:     | 73.64%   | Conservative:    |
| Best Local Similarity:  | 58.91%   | Mismatches:      |
| Query Match:            | 58.33%   | Indels:          |
| DB:                     | 7        | Gaps:            |
|                         |          | 1                |
| US-10-780-002-2 (1-260) | x        | CN190989 (1-865) |



**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: September 12, 2005, 17:23:43 ; Search time 482 Seconds  
(without alignments)  
3193.221 Million cell updates/sec

Title: US-10-780-002-2  
Perfect score: 1368  
Sequence: 1 MKEGKHFLVHGACGWSW.....MLCEPQKLCASLEIAHKYN 260

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+\_p2n.model -DEV=xlp  
-Q=/cgn2.1/USPTO.spool\_p/US10780002/runat\_11092005\_104746\_26588/app\_query.fasta\_1.455  
-DB=N Geneseq 16Dec04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10780002@cgn\_1.1.708 @runat\_11092005\_104746\_26588 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 16Dec04:.\*  
1: geneseqn1980s:.\*  
2: geneseqn1990s:.\*  
3: geneseqn2000s:.\*  
4: geneseqn2001as:.\*  
5: geneseqn2001bs:.\*  
6: geneseqn2002as:.\*  
7: geneseqn2002bs:.\*  
8: geneseqn2003as:.\*  
9: geneseqn2003bs:.\*  
10: geneseqn2003cs:.\*  
11: geneseqn2003ds:.\*  
12: geneseqn2004as:.\*  
13: geneseqn2004bs:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES  |       |       |        |                      |
|------------|-------|-------|--------|----------------------|
| Result No. | Score | Match | Length | Description          |
| 1          | 1368  | 100.0 | 1079   | 8 ADA26498 Tobacco s |
| 2          | 768   | 56.1  | 792    | 6 ABZ13515 Arabidops |
| 3          | 768   | 56.1  | 816    | 3 AAC42344 Arabidops |
| 4          | 768   | 56.1  | 918    | 8 ADA26530 Arabidops |
| 5          | 696   | 50.9  | 870    | 3 AAC34411 Arabidops |

|    |       |      |      |                       |
|----|-------|------|------|-----------------------|
| 6  | 695   | 50.8 | 792  | 6 ABZ13746 Arabidops  |
| 7  | 695   | 50.8 | 864  | 3 AAC47519 Arabidops  |
| 8  | 695   | 50.8 | 950  | 8 ADA26531 Arabidops  |
| 9  | 690   | 50.4 | 792  | 3 AAC43133 Arabidops  |
| 10 | 690   | 50.4 | 792  | 6 ABZ12412 Arabidops  |
| 11 | 690   | 50.4 | 792  | 12 ADN74620 Arabidops |
| 12 | 690   | 50.4 | 1152 | 8 ADA26532 Arabidops  |
| 13 | 680.5 | 49.7 | 1109 | 8 ADA26533 Arabidops  |
| 14 | 665.5 | 48.6 | 1046 | 8 ADA26535 Arabidops  |
| 15 | 656   | 48.0 | 984  | 3 AAC49239 Arabidops  |
| 16 | 654.5 | 47.8 | 834  | 8 ADA26537 Arabidops  |
| 17 | 645   | 47.1 | 834  | 8 ADA26536 Arabidops  |
| 18 | 640.5 | 46.8 | 770  | 8 ADA67827 Arabidops  |
| 19 | 640.5 | 46.8 | 771  | 6 ABZ13473 Arabidops  |
| 20 | 640.5 | 46.8 | 823  | 3 AAC42529 Arabidops  |
| 21 | 640.5 | 46.8 | 928  | 8 ADA26538 Arabidops  |
| 22 | 635   | 46.4 | 909  | 3 AAC47800 Arabidops  |
| 23 | 635   | 46.4 | 984  | 8 ADA26534 Arabidops  |
| 24 | 635   | 46.4 | 1056 | 3 AAC47895 Arabidops  |
| 25 | 632.5 | 46.2 | 795  | 8 ADA70991 Rice gene  |
| 26 | 603.5 | 44.1 | 1325 | 3 AAC47916 Arabidops  |
| 27 | 590   | 43.1 | 828  | 3 AAC41733 Arabidops  |
| 28 | 590   | 43.1 | 908  | 8 ADA26539 Arabidops  |
| 29 | 589   | 43.1 | 903  | 3 AAC34141 Arabidops  |
| 30 | 585   | 42.8 | 817  | 2 AAT86824 Hevea bra  |
| 31 | 585   | 42.8 | 1091 | 2 AAT36351 Hevea bra  |
| 32 | 574.5 | 42.0 | 978  | 8 ADA69955 Rice gene  |
| 33 | 548   | 40.1 | 860  | 8 ADA71305 Rice gene  |
| 34 | 530.5 | 38.8 | 777  | 12 ADQ16305 Nucleotid |
| 35 | 530.5 | 38.8 | 777  | 12 ADQ16306 Nucleotid |
| 36 | 501.5 | 36.7 | 783  | 8 ADA70766 Rice gene  |
| 37 | 438   | 32.0 | 720  | 8 ADA70729 Rice gene  |
| 38 | 426.5 | 31.2 | 422  | 3 AAC43356 Arabidops  |
| 39 | 393.5 | 28.8 | 1550 | 8 ADA26540 Arabidops  |
| 40 | 385.5 | 28.2 | 1155 | 10 ABT23286 Seed deve |
| 41 | 380   | 27.8 | 1815 | 3 AAC44497 Zea mays   |
| 42 | 377.5 | 27.6 | 1665 | 8 ADA26542 Arabidops  |
| 43 | 368   | 26.9 | 1268 | 3 AAC32766 Arabidops  |
| 44 | 368   | 26.9 | 1751 | 8 ADA26543 Arabidops  |
| 45 | 367   | 26.8 | 771  | 3 AAC43020 Arabidops  |

## ALIGNMENTS

RESULT 1  
ADA26498  
ID ADA26498 standard; cDNA; 1079 BP.  
XX  
AC ADA26498;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Tobacco salicylic acid binding protein 2 (SABP2) gene.  
XX  
KW salicylic acid-binding protein; SABP2; disease resistance; plant;  
KW modulator; tobacco; ds; gene.  
XX  
OS Nicotiana tabacum.  
XX  
FH Key Location/Qualifiers  
FT CDS 42..824  
FT /\*tag= a  
FT /product= "salicylic acid binding protein SABP2"  
XX  
PN WO2003016551-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 16-AUG-2002; 2002WO-US026312.  
PR 16-AUG-2001; 2001US-0312863P.  
XX  
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.



|  |             |   |     |
|--|-------------|---|-----|
| Best Local Similarity: 57.53%              |             | Matches: 73   |     |
| Query Match: 56.14%                        |             | Indels: 2   |     |
| DB: 6                                      |             | Gaps: 2   |     |
| US-10-780-002-2 (1-260) x AB213515 (1-792) |             |   |     |
| QY   | 2           | LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTyrTyr            | 21  |
| DB   | 13          | AAGAGGAACAACATTTTGTACTAGTACATGGTTCGTGCCATGGCGGTGGTCTGGTAC               | 72  |
| QY   | 22          | LysLeuLysProLeuLeuGluAlaLaGlyHisLysValThrAlaLeuAspLeuAlaAla             | 41  |
| DB   | 73          | AAGTTTAAGCGCTGTAGAGCGGTGGGCCACCGCTAACTGCTGTGGACTTAGTGCC                 | 132 |
| QY   | 42          | SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu            | 60  |
| DB   | 133         | TCCGGAATAGACAACAGAGGTGCATCTCGACATCCCCACATCGCAACAATCTCGGAG               | 192 |
| QY   | 61          | ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis            | 80  |
| DB   | 193         | CCATTGACGAAGCTCTGACCTCATTTGCCAATGATGAAGGTGTGCTCGTGTGTAC                 | 252 |
| QY   | 81          | SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla            | 100 |
| DB   | 253         | AGCTTTGGTGGCTTGAACCTTAGCCATAGCCATAGCAAGTTCGCCGAAAAATCTCTGC              | 312 |
| QY   | 101         | AlaValPheLeuAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu               | 120 |
| DB   | 313         | GCTGTATTCTTGACTCTTTCATGCGGACACCGAACCTCACCTTCCTCGTCTTGGAC                | 372 |
| QY   | 121         | GlnTyrAsnGluArgThrProAlaGluAsnTyrLeuAspThrGlnPheLeuProTyrGly            | 140 |
| DB   | 373         | AGTTTGGAGAACACATGCTCTCAAGAACATGATGGGACCGAATTCGAACCTTATGGT               | 432 |
| QY   | 141         | SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu            | 160 |
| DB   | 433         | TCAGACAATTCGGGACTG---AGTATGTTTTTTAGCCCTGACTTCATGAAGTTGGGTCTC            | 489 |
| QY   | 161         | TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer            | 180 |
| DB   | 490         | TACCAGCTTTCTCCAGTTGAGGATCTTGAACCTGGGATTACTTTTATGAGGCCAGGATCG            | 549 |
| QY   | 181         | LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal            | 200 |
| DB   | 550         | TTATTTATTAAACGATTATACGAAGATGAAAACCTTCTCGGATGAAGGATATGGTCTGTT            | 609 |
| QY   | 201         | LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTyr            | 220 |
| DB   | 610         | CCTCGAGTTTTCATAGTGTGTAAGAGACAAAGCAATTCAGAAAGAACGCCAGAGATGG              | 669 |
| QY   | 221         | GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla            | 240 |
| DB   | 670         | ATGATTGATAATTTCCGGTGAATTTAGTATGAGATGGAGGAGACAGATCATATGCCA               | 729 |
| QY   | 241         | MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr               | 259 |
| DB   | 730         | ATGTTCTGAAGCCTCAGCAACTCAGTGATTACTTCTGAAAATCGCGGACAAATTC                 | 786 |
| RESULT 3                                   |             |   |     |
| ID   | AAC42344    | standard; DNA; 816 BP.  |     |
| XX   | AC          | AAC42344;   |     |
| XX   | AC          | AAC42344;   |     |
| DT   | 17-OCT-2000 | (first entry)   |     |
| XX   | DE          | Arabidopsis thaliana DNA fragment SEQ ID NO: 35196.                     |     |
| XX   | DE          | Hybridisation assay; genetic mapping; gene expression control;          |     |
| KW   | KW          | protein identification; signal transduction pathway; metabolic pathway; |     |
| KW   | KW          | promoter; termination sequence; ss.                                     |     |
| XX   | OS          | Arabidopsis thaliana.   |     |

|  |              |  |
|--|--------------|--|
| PR   | 01-JUL-1999; | 99US-0141842P.   |
| PR   | 01-JUL-1999; | 99US-0142154P.   |
| PR   | 02-JUL-1999; | 99US-0142055P.   |
| PR   | 06-JUL-1999; | 99US-0142390P.   |
| PR   | 08-JUL-1999; | 99US-0142803P.   |
| PR   | 09-JUL-1999; | 99US-0142920P.   |
| PR   | 12-JUL-1999; | 99US-0142977P.   |
| PR   | 13-JUL-1999; | 99US-0143542P.   |
| PR   | 14-JUL-1999; | 99US-0143624P.   |
| PR   | 15-JUL-1999; | 99US-0144005P.   |
| PR   | 16-JUL-1999; | 99US-0144085P.   |
| PR   | 16-JUL-1999; | 99US-0144086P.   |
| PR   | 19-JUL-1999; | 99US-0144325P.   |
| PR   | 19-JUL-1999; | 99US-0144331P.   |
| PR   | 19-JUL-1999; | 99US-0144332P.   |
| PR   | 19-JUL-1999; | 99US-0144333P.   |
| PR   | 19-JUL-1999; | 99US-0144334P.   |
| PR   | 19-JUL-1999; | 99US-0144335P.   |
| PR   | 20-JUL-1999; | 99US-0144352P.   |
| PR   | 20-JUL-1999; | 99US-0144632P.   |
| PR   | 20-JUL-1999; | 99US-0144884P.   |
| PR   | 21-JUL-1999; | 99US-0144814P.   |
| PR   | 21-JUL-1999; | 99US-0145086P.   |
| PR   | 21-JUL-1999; | 99US-0145088P.   |
| PR   | 22-JUL-1999; | 99US-0145085P.   |
| PR   | 22-JUL-1999; | 99US-0145087P.   |
| PR   | 22-JUL-1999; | 99US-0145276P.   |
| PR   | 27-JUL-1999; | 99US-0145913P.   |
| PR   | 27-JUL-1999; | 99US-0145918P.   |
| PR   | 27-JUL-1999; | 99US-0145919P.   |
| PR   | 28-JUL-1999; | 99US-0145951P.   |
| PR   | 02-AUG-1999; | 99US-0146386P.   |
| PR   | 02-AUG-1999; | 99US-0146388P.   |
| PR   | 02-AUG-1999; | 99US-0146389P.   |
| PR   | 03-AUG-1999; | 99US-0147038P.   |
| PR   | 04-AUG-1999; | 99US-0147204P.   |
| PR   | 04-AUG-1999; | 99US-0147302P.   |
| PR   | 05-AUG-1999; | 99US-0147192P.   |
| PR   | 05-AUG-1999; | 99US-0147260P.   |
| PR   | 06-AUG-1999; | 99US-0147303P.   |
| PR   | 06-AUG-1999; | 99US-0147416P.   |
| PR   | 09-AUG-1999; | 99US-0147493P.   |
| PR   | 09-AUG-1999; | 99US-0147935P.   |
| PR   | 10-AUG-1999; | 99US-0148171P.   |
| PR   | 11-AUG-1999; | 99US-0148319P.   |
| PR   | 12-AUG-1999; | 99US-0148341P.   |
| PR   | 13-AUG-1999; | 99US-0148565P.   |
| PR   | 16-AUG-1999; | 99US-0149368P.   |
| PR   | 17-AUG-1999; | 99US-0149317SP.  |
| PR   | 18-AUG-1999; | 99US-0149426P.   |
| PR   | 20-AUG-1999; | 99US-0149722P.   |
| PR   | 20-AUG-1999; | 99US-0149723P.   |
| PR   | 20-AUG-1999; | 99US-0149929P.   |
| PR   | 23-AUG-1999; | 99US-0149902P.   |
| PR   | 23-AUG-1999; | 99US-0149930P.   |
| PR   | 25-AUG-1999; | 99US-0150566P.   |
| PR   | 26-AUG-1999; | 99US-0150884P.   |
| PR   | 27-AUG-1999; | 99US-0151065P.   |
| PR   | 27-AUG-1999; | 99US-0151066P.   |
| PR   | 27-AUG-1999; | 99US-0151069P.   |
| PR   | 30-AUG-1999; | 99US-0151303P.   |
| PR   | 31-AUG-1999; | 99US-0151438P.   |
| PR   | 01-SEP-1999; | 99US-0151930P.   |
| PR   | 07-SEP-1999; | 99US-0152363P.   |
| PR   | 10-SEP-1999; | 99US-0153070P.   |
| PR   | 13-SEP-1999; | 99US-0153758P.   |
| PR   | 15-SEP-1999; | 99US-0154018P.   |
| PR   | 16-SEP-1999; | 99US-0154039P.   |
| PR   | 20-SEP-1999; | 99US-0154779P.   |
| PR   | 22-SEP-1999; | 99US-0155139P.   |
| PR   | 23-SEP-1999; | 99US-0155486P.   |
| PR   | 24-SEP-1999; | 99US-0155659P.   |
| PR   | 28-SEP-1999; | 99US-0156458P.   |
| PR   | 29-SEP-1999; | 99US-0156596P.   |
| PR   | 04-OCT-1999; | 99US-0157117P.   |
| PR   | 05-OCT-1999; | 99US-0157753P.   |
| PR   | 06-OCT-1999; | 99US-0157865P.   |
| PR   | 07-OCT-1999; | 99US-0158029P.   |
| PR   | 08-OCT-1999; | 99US-0158232P.   |
| PR   | 12-OCT-1999; | 99US-0158369P.   |
| PR   | 13-OCT-1999; | 99US-0159293P.   |
| PR   | 13-OCT-1999; | 99US-0159294P.   |
| PR   | 13-OCT-1999; | 99US-0159295P.   |
| PR   | 14-OCT-1999; | 99US-0159329P.   |
| PR   | 14-OCT-1999; | 99US-0159330P.   |
| PR   | 14-OCT-1999; | 99US-0159331P.   |
| PR   | 14-OCT-1999; | 99US-0159637P.   |
| PR   | 14-OCT-1999; | 99US-0159638P.   |
| PR   | 18-OCT-1999; | 99US-0159584P.   |
| PR   | 21-OCT-1999; | 99US-0160741P.   |
| PR   | 21-OCT-1999; | 99US-0160767P.   |
| PR   | 21-OCT-1999; | 99US-0160768P.   |
| PR   | 21-OCT-1999; | 99US-0160770P.   |
| PR   | 21-OCT-1999; | 99US-0160814P.   |
| PR   | 21-OCT-1999; | 99US-0160815P.   |
| PR   | 22-OCT-1999; | 99US-0160980P.   |
| PR   | 22-OCT-1999; | 99US-0160981P.   |
| PR   | 22-OCT-1999; | 99US-0160989P.   |
| PR   | 25-OCT-1999; | 99US-0161404P.   |
| PR   | 25-OCT-1999; | 99US-0161405P.   |
| PR   | 25-OCT-1999; | 99US-0161406P.   |
| PR   | 26-OCT-1999; | 99US-0161359P.   |
| PR   | 26-OCT-1999; | 99US-0161360P.   |
| PR   | 26-OCT-1999; | 99US-0161361P.   |
| PR   | 28-OCT-1999; | 99US-0161920P.   |
| PR   | 28-OCT-1999; | 99US-0161992P.   |
| PR   | 28-OCT-1999; | 99US-0161993P.   |
| PR   | 29-OCT-1999; | 99US-0162142P.   |
| Alignment Scores:                          |              |  |
| Pred. No.:                                 | 1.89e-79     | 816  |
| Score:                                     | 768.00       | 149  |
| Percent Similarity:                        | 71.04%       | Conservative: 35   |
| Best Local Similarity:                     | 57.53%       | Mismatches: 73   |
| Query Match:                               | 56.14%       | Indels: 2  |
| DB:  | 3            | Gaps: 2  |
| US-10-780-002-2 (1-260) x AAC42344 (1-816) |              |  |
| QY   | 2            | LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyr 21  |
| Db   | 37           | AAGAGGAACAACATTTGTACTAGTACATGTTGCGCCATGGCGCGTGGTGCTGCTAC 96      |
| QY   | 22           | LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla 41     |
| Db   | 97           | AAGTTTAAGCGCTGCTAGAGCGGTGGCCCGCTAACTGCTGTGGACTTAGCTGCC 156       |
| QY   | 42           | SerGlyThrAsp---LeuAtqLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu 60  |
| Db   | 157          | TCCGGATAGACACAACGAGGTGCGATCACTGACATCCCCCATCGGAACAATACTCGGAG 216  |
| QY   | 61           | ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80  |
| Db   | 217          | CCATTGACGAAGCTCCTGACCTCATTTGCCAAATGATGAAAGGTTGTGCTCGTTGGTCAC 276 |
| QY   | 81           | SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100 |
| Db   | 277          | AGCTTTGGTGGCTTGAACTTAGCCATGCCATGAAAGTTTCCCGAAAAAATCTCTGTC 336    |
| QY   | 101          | AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerPheValLeuGlu 120    |

Db 337 GCTGTATCTTCACTGCTTTCATGCGGACACCAACTCCACCATCTCTGCTCTGGAC 396  
 Qy 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrClnPheLeuProTyrGly 140  
 Db 397 AAGTTTGGAGAACATGCTCAAGAAAGCATGGATGGCAGCAATTCGAACCTTATGGT 456  
 Qy 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160  
 Db 457 TCAGACAAATCCGGACTG---AGTATGTTTTTTAGCCCTGACTTCATGAAGTTGGTCTC 513  
 Qy 161 TyrGlnLeuCysSerProGluLeuPheLeuAlaLeuAlaSerSerLeuValArgProSerSer 180  
 Db 514 TACCAGCTTCTCCAGTTCAGGACTTGAACCTGGGATTACTTTTATGAGGCGCAGGATCG 573  
 Qy 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200  
 Db 574 TTATTTTAAACAGATTTATCGAAGATGAAAACTTCTCGGATGAAGGATATGGGTCTGT 633  
 Qy 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220  
 Db 634 CCTCGAGTTTTCATAGTGTGTAAGAGGACAAAGCAATTCAGAGAACGCCAGAGATGG 693  
 Qy 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaPheHisMetAla 240  
 Db 694 ATGATTTGATAAATTTCCGGTGAATTTAGTGATGGAGATGGAGAGACAGATCATATGCCA 753  
 Qy 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259  
 Db 754 ATGTTCTGCAAGCCTCAGCAACTCAGTGATTAATCTCTGANAATCGCGACAATTC 810

## RESULT 4

ID ADA26530 standard; DNA; 918 BP.  
 XX ADA26530;  
 AC ADA26530;  
 DT 20-NOV-2003 (first entry)  
 XX Arabidopsis thaliana AtSB2L1 gene homologous to tobacco SABP2 gene.  
 XX salicylic acid-binding protein; SABP2; disease resistance; plant;  
 KW modulator; tobacco; ds; gene.  
 XX Arabidopsis thaliana.  
 XX WO2003016551-A2.  
 PN 27-FEB-2003.  
 XX 16-AUG-2002; 2002WO-US026312.  
 PF 16-AUG-2001; 2001US-0312863P.  
 PR (BOYC-) BOYCE THOMPSON INST PLANT RES INC.  
 PA Klessig DF, Kumar D;  
 PI WPI; 2003-278578/27.  
 DR

XX New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for  
 PT producing transgenic plants with increased resistance to disease, or for  
 PT screening SABP2 modulators that confer enhanced resistance of plants to  
 PT disease.

XX Claim 65; Page: 98pp; English.

XX The invention relates to a novel isolated 1079 bp salicylic acid-binding  
 CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid  
 CC protein, sequences that specifically hybridize with the nucleic acid, the  
 CC complement of the nucleic acid or a natural allelic variant of the  
 CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for  
 CC conferring increased resistance to disease in plants, or for screening

CC modulators of SABP2, which confer increased or enhanced resistance of  
 CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also  
 CC useful for producing transgenic plants with increased resistance to  
 CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful  
 CC for studying the molecular mechanisms responsible for the modulation of  
 CC salicylic acid-mediated disease resistance in plants. This sequence  
 CC represents the gene encoding the Arabidopsis thaliana SABP2 AtSB2L1 which  
 CC has homology to the gene (ADA26498) encoding the tobacco SABP2 protein.  
 CC (Note: this sequence is not given in the specification but is derived  
 CC from the Genbank accession number given by the inventors).

XX Sequence 918 BP; 266 A; 175 C; 221 G; 256 T; 0 U; 0 Other;

## Alignment Scores:

| Pred. No.:             | 2,23e-79 | Length:       | 918 |
|------------------------|----------|---------------|-----|
| Score:                 | 768.00   | Matches:      | 149 |
| Percent Similarity:    | 71.04%   | Conservative: | 35  |
| Best Local Similarity: | 57.53%   | Mismatches:   | 73  |
| Query Match:           | 56.14%   | Indels:       | 2   |
| DB:                    | 8        | Gaps:         | 2   |

US-10-780-002-2 (1-260) x ADA26530 (1-918)

Qy 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyTyrSerTrpTyr 21  
 Db 56 AAGAGGAAACAACATTTTGTACTAGTACATGTTCTGCCATGGCGCTGGTGTGATC 115  
 Qy 22 LysLeuLysProLeuLeuGluAlaGlyHisValThrAlaLeuAspLeuAla 41  
 Db 116 AAGTTTAAAGCCGCTGCTAGAGCGGTGGGCCACCGCTAACTGCTGTGACTAGTCC 175  
 Qy 42 SerGlyThrAsp---LeuArgLysIleGluLeuLeuArgThrLeuTyrAspTyrThrLeu 60  
 Db 176 TCCGGAATAGACACACGAGGTTCGATCACTGCATCCCCACATGCGAACATATCTCGGAG 235  
 Qy 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80  
 Db 236 CCATTGACGAAGCTCCTGACCTCATTTGCCAAATGATGAAAGGTTGTCTGCTGGTCCAC 295  
 Qy 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100  
 Db 296 AGCTTTGGTGGCTTGAACCTTAGCCATGCCATGAAAGTTTCCCGAAAAATCTCTGTC 355  
 Qy 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120  
 Db 356 GCTGTATTTCTTGACTGCTTTTCATGCCGACACCCGACACTCACCATCTCTGCTTGAC 415  
 Qy 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140  
 Db 416 AAGTTTGGAGCAACATGCTCAAGAGCATGGATGGCCACCGAATTCGAACCTTATGGT 475  
 Qy 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160  
 Db 476 TCAGACAAATCCGGACTG---AGTATGTTTTTTAGCCCTGACTTCATGAAGTTGGGTCTC 532  
 Qy 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180  
 Db 533 TACCAGCTTCTCCAGTTGAGGATCTTGAACCTGGGATTACTTTTATGAGGCCAGGATCG 592  
 Qy 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200  
 Db 593 TTATTTTAAACAGATTTATCGAAGATGAAAACTTCTCGGATGAAGGATATGGGTCTGTT 652  
 Qy 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220  
 Db 653 CCTCGAGTTTTCATAGTGTGTAAGAGGACAAAGCAATTCAGAGAAACGCCAGAGATGG 712  
 Qy 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240  
 Db 713 ATGATTTGATAAATTTCCGGTGAATTTAGTGATGGAGATGGAGGACAGATCATATGCCA 772  
 Qy 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259

|           |   |  |                |
|-----------|---|--|----------------|
| Db        | 773   | ATGTTCTGCAGCCTCAGCAACTCAGTGATTACTTCTGAAATCGCGGACAAATTC | 829            |
| RESULT 5  |   |  |                |
| AAAC34411 |   |  |                |
| ID        | AAAC34411   | standard; DNA; 870 BP.                                 |                |
| XX        |   |  |                |
| AC        | AAAC34411;  |  |                |
| XX        |   |  |                |
| DT        | 17-OCT-2000   | (first entry)  |                |
| XX        |   |  |                |
| DE        | Arabidopsis thaliana  | DNA fragment SEQ ID NO: 6557.                          |                |
| XX        |   |  |                |
| KW        | Hybridisation assay; genetic mapping; gene expression control;          |  |                |
| KW        | protein identification; signal transduction pathway; metabolic pathway; |  |                |
| KW        | promoter; termination sequence; ss.                                     |  |                |
| XX        |   |  |                |
| OS        | Arabidopsis thaliana.   |  |                |
| XX        |   |  |                |
| PN        | EP1033405-A2.   |  |                |
| XX        |   |  |                |
| PD        | 06-SEP-2000.  |  |                |
| XX        |   |  |                |
| PF        | 25-FEB-2000; 2000EP-00301439.   |  |                |
| XX        |   |  |                |
| PR        | 25-FEB-1999;  | 99US-0121825P.   | 99US-0139457P. |
| PR        | 05-MAR-1999;  | 99US-0123180P.   | 99US-0139458P. |
| PR        | 09-MAR-1999;  | 99US-0123548P.   | 99US-0139459P. |
| PR        | 23-MAR-1999;  | 99US-0125788P.   | 99US-0139460P. |
| PR        | 25-MAR-1999;  | 99US-0126264P.   | 99US-0139461P. |
| PR        | 29-MAR-1999;  | 99US-0126785P.   | 99US-0139462P. |
| PR        | 01-APR-1999;  | 99US-0127462P.   | 99US-0139463P. |
| PR        | 06-APR-1999;  | 99US-0128234P.   | 99US-0139750P. |
| PR        | 08-APR-1999;  | 99US-0128714P.   | 99US-0139763P. |
| PR        | 16-APR-1999;  | 99US-0129845P.   | 99US-0139817P. |
| PR        | 19-APR-1999;  | 99US-0130077P.   | 99US-0139899P. |
| PR        | 21-APR-1999;  | 99US-0130449P.   | 99US-0140353P. |
| PR        | 23-APR-1999;  | 99US-0130510P.   | 99US-0140354P. |
| PR        | 28-APR-1999;  | 99US-0131449P.   | 99US-0140355P. |
| PR        | 30-APR-1999;  | 99US-0132048P.   | 99US-0140695P. |
| PR        | 30-APR-1999;  | 99US-0132407P.   | 99US-0140823P. |
| PR        | 04-MAY-1999;  | 99US-0132484P.   | 99US-0140991P. |
| PR        | 05-MAY-1999;  | 99US-0132485P.   | 99US-0141287P. |
| PR        | 06-MAY-1999;  | 99US-0132486P.   | 99US-0141842P. |
| PR        | 06-MAY-1999;  | 99US-0132487P.   | 99US-0142154P. |
| PR        | 11-MAY-1999;  | 99US-0132863P.   | 99US-0142055P. |
| PR        | 14-MAY-1999;  | 99US-0134218P.   | 99US-0142390P. |
| PR        | 14-MAY-1999;  | 99US-0134219P.   | 99US-0142803P. |
| PR        | 14-MAY-1999;  | 99US-0134221P.   | 99US-0142977P. |
| PR        | 14-MAY-1999;  | 99US-0134370P.   | 99US-0143542P. |
| PR        | 18-MAY-1999;  | 99US-0134768P.   | 99US-0143624P. |
| PR        | 19-MAY-1999;  | 99US-0134941P.   | 99US-0143625P. |
| PR        | 20-MAY-1999;  | 99US-0135124P.   | 99US-0144086P. |
| PR        | 21-MAY-1999;  | 99US-0135353P.   | 99US-0144325P. |
| PR        | 24-MAY-1999;  | 99US-0135629P.   | 99US-0144331P. |
| PR        | 25-MAY-1999;  | 99US-0136021P.   | 99US-0144332P. |
| PR        | 27-MAY-1999;  | 99US-0136392P.   | 99US-0144333P. |
| PR        | 28-MAY-1999;  | 99US-0136782P.   | 99US-0144334P. |
| PR        | 01-JUN-1999;  | 99US-0137222P.   | 99US-0144352P. |
| PR        | 03-JUN-1999;  | 99US-0137528P.   | 99US-0144632P. |
| PR        | 04-JUN-1999;  | 99US-0137502P.   | 99US-0144884P. |
| PR        | 07-JUN-1999;  | 99US-0137724P.   | 99US-0145086P. |
| PR        | 08-JUN-1999;  | 99US-0138094P.   | 99US-0145088P. |
| PR        | 10-JUN-1999;  | 99US-0138540P.   | 99US-0145085P. |
| PR        | 10-JUN-1999;  | 99US-0138847P.   | 99US-0145087P. |
| PR        | 14-JUN-1999;  | 99US-0139119P.   | 99US-0145089P. |
| PR        | 16-JUN-1999;  | 99US-0139452P.   | 99US-0145132P. |
| PR        | 16-JUN-1999;  | 99US-0139453P.   | 99US-0145145P. |
| PR        | 17-JUN-1999;  | 99US-0139492P.   | 99US-0145218P. |
| PR        | 18-JUN-1999;  | 99US-0139454P.   | 99US-0145224P. |
| PR        | 18-JUN-1999;  | 99US-0139455P.   | 99US-0145276P. |
| PR        | 18-JUN-1999;  | 99US-0139456P.   | 99US-0145913P. |
| PR        | 18-JUN-1999;  | 99US-0139457P.   | 99US-0145918P. |
| PR        | 18-JUN-1999;  | 99US-0139458P.   | 99US-0145919P. |
| PR        | 18-JUN-1999;  | 99US-0139459P.   | 99US-0145951P. |
| PR        | 18-JUN-1999;  | 99US-0139460P.   | 99US-0146386P. |
| PR        | 18-JUN-1999;  | 99US-0139461P.   | 99US-0146388P. |
| PR        | 18-JUN-1999;  | 99US-0139462P.   | 99US-0146389P. |
| PR        | 18-JUN-1999;  | 99US-0139463P.   | 99US-0147038P. |
| PR        | 18-JUN-1999;  | 99US-0139750P.   | 99US-0147204P. |
| PR        | 18-JUN-1999;  | 99US-0139763P.   | 99US-0147302P. |
| PR        | 18-JUN-1999;  | 99US-0139817P.   | 99US-0147303P. |
| PR        | 18-JUN-1999;  | 99US-0139899P.   | 99US-0147260P. |
| PR        | 18-JUN-1999;  | 99US-0140353P.   | 99US-0147303P. |
| PR        | 18-JUN-1999;  | 99US-0140354P.   | 99US-0147453P. |
| PR        | 18-JUN-1999;  | 99US-0140355P.   | 99US-0147454P. |
| PR        | 18-JUN-1999;  | 99US-0140695P.   | 99US-0147935P. |
| PR        | 18-JUN-1999;  | 99US-0140823P.   | 99US-0148171P. |
| PR        | 18-JUN-1999;  | 99US-0140991P.   | 99US-0148171P. |
| PR        | 18-JUN-1999;  | 99US-0141287P.   | 99US-0148319P. |
| PR        | 18-JUN-1999;  | 99US-0141842P.   | 99US-0148341P. |
| PR        | 18-JUN-1999;  | 99US-0142154P.   | 99US-0148565P. |
| PR        | 18-JUN-1999;  | 99US-0142055P.   | 99US-0148684P. |
| PR        | 18-JUN-1999;  | 99US-0142390P.   | 99US-0149368P. |
| PR        | 18-JUN-1999;  | 99US-0142803P.   | 99US-0149175P. |
| PR        | 18-JUN-1999;  | 99US-0142977P.   | 99US-0149455P. |
| PR        | 18-JUN-1999;  | 99US-0143542P.   | 99US-0149456P. |
| PR        | 18-JUN-1999;  | 99US-0143624P.   |                |
| PR        | 18-JUN-1999;  | 99US-0143625P.   |                |
| PR        | 18-JUN-1999;  | 99US-0144086P.   |                |
| PR        | 18-JUN-1999;  | 99US-0144325P.   |                |
| PR        | 18-JUN-1999;  | 99US-0144331P.   |                |
| PR        | 18-JUN-1999;  | 99US-0144332P.   |                |
| PR        | 18-JUN-1999;  | 99US-0144333P.   |                |
| PR        | 18-JUN-1999;  | 99US-0144334P.   |                |
| PR        | 18-JUN-1999;  | 99US-0144352P.   |                |
| PR        | 18-JUN-1999;  | 99US-0144632P.   |                |
| PR        | 18-JUN-1999;  | 99US-0144884P.   |                |
| PR        | 18-JUN-1999;  | 99US-0145086P.   |                |
| PR        | 18-JUN-1999;  | 99US-0145088P.   |                |
| PR        | 18-JUN-1999;  | 99US-0145085P.   |                |
| PR        | 18-JUN-1999;  | 99US-0145087P.   |                |
| PR        | 18-JUN-1999;  | 99US-0145089P.   |                |
| PR        | 18-JUN-1999;  | 99US-0145132P.   |                |
| PR        | 18-JUN-1999;  | 99US-0145145P.   |                |
| PR        | 18-JUN-1999;  | 99US-0145218P.   |                |
| PR        | 18-JUN-1999;  | 99US-0145224P.   |                |
| PR        | 18-JUN-1999;  | 99US-0145276P.   |                |
| PR        | 18-JUN-1999;  | 99US-0145913P.   |                |
| PR        | 18-JUN-1999;  | 99US-0145918P.   |                |
| PR        | 18-JUN-1999;  | 99US-0145919P.   |                |
| PR        | 18-JUN-1999;  | 99US-0145951P.   |                |
| PR        | 18-JUN-1999;  | 99US-0146386P.   |                |
| PR        | 18-JUN-1999;  | 99US-0146388P.   |                |
| PR        | 18-JUN-1999;  | 99US-0146389P.   |                |
| PR        | 18-JUN-1999;  | 99US-0147038P.   |                |
| PR        | 18-JUN-1999;  | 99US-0147204P.   |                |
| PR        | 18-JUN-1999;  | 99US-0147302P.   |                |
| PR        | 18-JUN-1999;  | 99US-0147303P.   |                |
| PR        | 18-JUN-1999;  | 99US-0147260P.   |                |
| PR        | 18-JUN-1999;  | 99US-0147303P.   |                |
| PR        | 18-JUN-1999;  | 99US-0147453P.   |                |
| PR        | 18-JUN-1999;  | 99US-0147454P.   |                |
| PR        | 18-JUN-1999;  | 99US-0147935P.   |                |
| PR        | 18-JUN-1999;  | 99US-0148171P.   |                |
| PR        | 18-JUN-1999;  | 99US-0148171P.   |                |
| PR        | 18-JUN-1999;  | 99US-0148319P.   |                |
| PR        | 18-JUN-1999;  | 99US-0148341P.   |                |
| PR        | 18-JUN-1999;  | 99US-0148565P.   |                |
| PR        | 18-JUN-1999;  | 99US-0148684P.   |                |
| PR        | 18-JUN-1999;  | 99US-0149368P.   |                |
| PR        | 18-JUN-1999;  | 99US-0149175P.   |                |
| PR        | 18-JUN-1999;  | 99US-0149455P.   |                |
| PR        | 18-JUN-1999;  | 99US-0149456P.   |                |



PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
PI Harper JF, Kreps J, Wang X, Zhu T;  
XX  
XX WPI; 2002-304127/34.  
XX  
PT Identifying a stress condition to which a plant cell has been exposed and  
PT producing plants with increased tolerance to these abiotic stresses.  
XX  
XX Claim 144; SEQ ID NO 1551; 577pp + Sequence Listing; English.  
XX  
CC The invention relates to identifying a stress condition to which a plant  
CC cell has been exposed, comprising: (a) contacting nucleic acid to which a plant  
CC representative of expressed polynucleotides in the plant cell with an  
CC array or probes representative of the plant cell genome; and (b)  
CC detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present sequence is that  
CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used  
CC in methods of the invention. Note: The sequence data for this patent is  
CC not represented in the printed specification but is based on sequence  
CC information supplied to Derwent by the European Patent Office  
XX  
SQ Sequence 792 BP; 210 A; 169 C; 187 G; 226 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 5,81e-71 Length: 792  
Score: 695.00 Matches: 138  
Percent Similarity: 67.18% Conservative: 36  
Best Local Similarity: 53.28% Mismatches: 83  
Query Match: 50.80% Indels: 2  
DB: 6 Gaps: 2  
  
US-10-780-002-2 (1-260) x AB213746 (1-792)  
  
QY 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyr 21  
DB 13 AAGAGGAACCAACACTTCGTGTAGTACATGCTGGTCCGCGCATGGTGGTGGTAC 72  
  
QY 22 LysLeuLysProLeuLeuAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAla 41  
DB 73 AAGTTAAGCCTTCTCGAGGCTTGGCCATCGTGTAACCGCTTAGACCTAGCTGCT 132  
  
QY 42 SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu 60  
DB 133 TCCGGTATAGACACCAACAGGTCATCTACATGCTTCTACATGGAACAATATCTGAG 192  
  
QY 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80  
DB 193 CCATTGATGCGCTAATGACTTCATTGCCGAATGATGAGAGGTTGACTCGTGGTCAAT 252  
  
QY 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100  
DB 253 AGCTTTGAGGTTTGAGTTAGCTTAGCCATGATGATAGTTCCCGATAAATCTCTGTC 312  
  
QY 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120  
DB 313 TCTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 372  
  
QY 121 GlnTyrAsnGluArgThrProAlaGluAsnTyrLeuAspThrGlnPheLeuProTyrGly 140  
DB 373 AAGTTTGCAGGAGCATGACACAGAGGATGGATGGGCTCTGAGCTCGAGACATATGCT 432  
  
QY 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160  
DB 433 TCAGATAATTCGGCTTG---TCTGTGTCTTTCAGCAGCGACTTCATGAAGCAGCTCTC 489  
  
QY 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180  
DB 490 TACCAACTTCTCCTCTGAGGATCTTGAAGCTTGGATTTGCTTAAAGAGGCGCTAGTTCA 549  
  
QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200  
  
Db 550 TTTGTTTATTAATGAATATCGAATCGAGAACCTTTCTGAGAAAGGATGATCTGTT 609  
QY 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220  
Db 610 CCTGAGCTTACATTGTGTGCAAGAGGACAAACATTATCTCGAAGACCATCAACGATGG 669  
QY 221 GlnIleAspAniIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240  
Db 670 ATGATCCATAATATCCGCGCAATTTAGTATTGAGATGGAAGAGACTGATCATATGCCA 729  
QY 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuAlaHisLysTyr 259  
Db 730 ATGTTTGCACAACTCACTACTAGTGACCATCTATTGGCAATCGCTGACAAATTC 786  
  
RESULT 7  
AAC47519  
ID AAC47519 standard; DNA; 864 BP.  
XX AAC47519;  
AC AAC47519;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 54118.  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0127462P.  
PR 01-APR-1999; 99US-0128234P.  
PR 06-APR-1999; 99US-0128714P.  
PR 08-APR-1999; 99US-0129845P.  
PR 16-APR-1999; 99US-0130077P.  
PR 19-APR-1999; 99US-0130449P.  
PR 21-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.



|                 |                 |                 |                |
|-----------------|-----------------|-----------------|----------------|
| PR 03-JUN-1999; | 99US-01337528P. | PR 05-AUG-1999; | 99US-0147260P. |
| PR 04-JUN-1999; | 99US-0137502P.  | PR 06-AUG-1999; | 99US-0147303P. |
| PR 07-JUN-1999; | 99US-0137724P.  | PR 06-AUG-1999; | 99US-0147416P. |
| PR 08-JUN-1999; | 99US-0138094P.  | PR 09-AUG-1999; | 99US-0147493P. |
| PR 10-JUN-1999; | 99US-0138540P.  | PR 09-AUG-1999; | 99US-0147935P. |
| PR 10-JUN-1999; | 99US-0138847P.  | PR 10-AUG-1999; | 99US-0148171P. |
| PR 14-JUN-1999; | 99US-0139119P.  | PR 11-AUG-1999; | 99US-0148319P. |
| PR 16-JUN-1999; | 99US-0139452P.  | PR 12-AUG-1999; | 99US-0148341P. |
| PR 16-JUN-1999; | 99US-0139453P.  | PR 13-AUG-1999; | 99US-0148565P. |
| PR 17-JUN-1999; | 99US-0139492P.  | PR 13-AUG-1999; | 99US-0148684P. |
| PR 18-JUN-1999; | 99US-0139454P.  | PR 16-AUG-1999; | 99US-0149368P. |
| PR 18-JUN-1999; | 99US-0139455P.  | PR 17-AUG-1999; | 99US-0149175P. |
| PR 18-JUN-1999; | 99US-0139456P.  | PR 18-AUG-1999; | 99US-0149426P. |
| PR 18-JUN-1999; | 99US-0139457P.  | PR 20-AUG-1999; | 99US-0149723P. |
| PR 18-JUN-1999; | 99US-0139458P.  | PR 20-AUG-1999; | 99US-0149723P. |
| PR 18-JUN-1999; | 99US-0139459P.  | PR 20-AUG-1999; | 99US-0149929P. |
| PR 18-JUN-1999; | 99US-0139460P.  | PR 23-AUG-1999; | 99US-0149902P. |
| PR 18-JUN-1999; | 99US-0139461P.  | PR 23-AUG-1999; | 99US-0149930P. |
| PR 18-JUN-1999; | 99US-0139462P.  | PR 25-AUG-1999; | 99US-0150566P. |
| PR 18-JUN-1999; | 99US-0139463P.  | PR 26-AUG-1999; | 99US-0150884P. |
| PR 18-JUN-1999; | 99US-0139750P.  | PR 27-AUG-1999; | 99US-0151065P. |
| PR 18-JUN-1999; | 99US-0139763P.  | PR 27-AUG-1999; | 99US-0151066P. |
| PR 21-JUN-1999; | 99US-0139817P.  | PR 27-AUG-1999; | 99US-0151080P. |
| PR 22-JUN-1999; | 99US-0139899P.  | PR 30-AUG-1999; | 99US-0151303P. |
| PR 23-JUN-1999; | 99US-0140353P.  | PR 31-AUG-1999; | 99US-0151438P. |
| PR 23-JUN-1999; | 99US-0140354P.  | PR 01-SEP-1999; | 99US-0151930P. |
| PR 24-JUN-1999; | 99US-0140695P.  | PR 07-SEP-1999; | 99US-0152363P. |
| PR 28-JUN-1999; | 99US-0140823P.  | PR 10-SEP-1999; | 99US-0153070P. |
| PR 29-JUN-1999; | 99US-0140991P.  | PR 13-SEP-1999; | 99US-0153758P. |
| PR 30-JUN-1999; | 99US-0141128P.  | PR 15-SEP-1999; | 99US-0154018P. |
| PR 01-JUL-1999; | 99US-0141184P.  | PR 16-SEP-1999; | 99US-0154039P. |
| PR 01-JUL-1999; | 99US-0142154P.  | PR 20-SEP-1999; | 99US-0154779P. |
| PR 02-JUL-1999; | 99US-0142055P.  | PR 22-SEP-1999; | 99US-0155139P. |
| PR 06-JUL-1999; | 99US-0142390P.  | PR 23-SEP-1999; | 99US-0155486P. |
| PR 08-JUL-1999; | 99US-0142803P.  | PR 24-SEP-1999; | 99US-0155659P. |
| PR 09-JUL-1999; | 99US-0142920P.  | PR 28-SEP-1999; | 99US-0156458P. |
| PR 12-JUL-1999; | 99US-0142977P.  | PR 29-SEP-1999; | 99US-0156596P. |
| PR 13-JUL-1999; | 99US-0143542P.  | PR 04-OCT-1999; | 99US-0157117P. |
| PR 14-JUL-1999; | 99US-0143624P.  | PR 05-OCT-1999; | 99US-0157753P. |
| PR 15-JUL-1999; | 99US-0144005P.  | PR 06-OCT-1999; | 99US-0157865P. |
| PR 16-JUL-1999; | 99US-0144085P.  | PR 07-OCT-1999; | 99US-0158029P. |
| PR 16-JUL-1999; | 99US-0144086P.  | PR 08-OCT-1999; | 99US-0158232P. |
| PR 19-JUL-1999; | 99US-0144325P.  | PR 12-OCT-1999; | 99US-0158369P. |
| PR 19-JUL-1999; | 99US-0144331P.  | PR 13-OCT-1999; | 99US-0159293P. |
| PR 19-JUL-1999; | 99US-0144332P.  | PR 13-OCT-1999; | 99US-0159294P. |
| PR 19-JUL-1999; | 99US-0144333P.  | PR 13-OCT-1999; | 99US-0159295P. |
| PR 19-JUL-1999; | 99US-0144334P.  | PR 14-OCT-1999; | 99US-0159329P. |
| PR 19-JUL-1999; | 99US-0144335P.  | PR 14-OCT-1999; | 99US-0159330P. |
| PR 20-JUL-1999; | 99US-0144352P.  | PR 14-OCT-1999; | 99US-0159331P. |
| PR 20-JUL-1999; | 99US-0144632P.  | PR 14-OCT-1999; | 99US-0159637P. |
| PR 21-JUL-1999; | 99US-0144884P.  | PR 14-OCT-1999; | 99US-0159638P. |
| PR 21-JUL-1999; | 99US-0144814P.  | PR 18-OCT-1999; | 99US-0159584P. |
| PR 21-JUL-1999; | 99US-0145086P.  | PR 21-OCT-1999; | 99US-0160741P. |
| PR 21-JUL-1999; | 99US-0145088P.  | PR 21-OCT-1999; | 99US-0160767P. |
| PR 22-JUL-1999; | 99US-0145085P.  | PR 21-OCT-1999; | 99US-0160768P. |
| PR 22-JUL-1999; | 99US-0145087P.  | PR 21-OCT-1999; | 99US-0160770P. |
| PR 22-JUL-1999; | 99US-0145089P.  | PR 21-OCT-1999; | 99US-0160814P. |
| PR 22-JUL-1999; | 99US-0145192P.  | PR 21-OCT-1999; | 99US-0160815P. |
| PR 23-JUL-1999; | 99US-0145145P.  | PR 22-OCT-1999; | 99US-0160980P. |
| PR 23-JUL-1999; | 99US-0145218P.  | PR 22-OCT-1999; | 99US-0160981P. |
| PR 23-JUL-1999; | 99US-0145224P.  | PR 22-OCT-1999; | 99US-0160989P. |
| PR 26-JUL-1999; | 99US-0145276P.  | PR 25-OCT-1999; | 99US-0161404P. |
| PR 27-JUL-1999; | 99US-0145913P.  | PR 25-OCT-1999; | 99US-0161405P. |
| PR 27-JUL-1999; | 99US-0145918P.  | PR 25-OCT-1999; | 99US-0161406P. |
| PR 28-JUL-1999; | 99US-0145919P.  | PR 26-OCT-1999; | 99US-0161359P. |
| PR 28-JUL-1999; | 99US-0145951P.  | PR 26-OCT-1999; | 99US-0161360P. |
| PR 02-AUG-1999; | 99US-0146386P.  | PR 26-OCT-1999; | 99US-0161361P. |
| PR 02-AUG-1999; | 99US-0146388P.  | PR 28-OCT-1999; | 99US-0161920P. |
| PR 03-AUG-1999; | 99US-0147038P.  | PR 28-OCT-1999; | 99US-0161992P. |
| PR 04-AUG-1999; | 99US-0147204P.  | PR 28-OCT-1999; | 99US-0161993P. |
| PR 05-AUG-1999; | 99US-0147302P.  | PR 29-OCT-1999; | 99US-0162142P. |
| PR 05-AUG-1999; | 99US-0147192P.  |                 |                |

Alignment Scores:

```
Pred. No.: 6,58e-71 Length: 864
Score: 695.00 Matches: 138
Percent Similarity: 67.18% Conservative: 36
Best Local Similarity: 53.28% Mismatches: 83
Query Match: 50.80% Indels: 2
DB: 3 Gaps: 2

US-10-780-002-2 (1-260) x AAG47519 (1-864)

QY 2 LysGluGlyLysPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyr 21
DB 31 AAGAGGAGCAACACTTCGTGTAGTACATGCTGCGGCCAGCGCATGGTGGTGTAC 90

QY 22 LysLeuLysProLeuLeuGluAlaGlyHisLysValThrAlaLeuAspLeuAla 41
DB 91 AAGGTTAAGCCCTCTTCGAGGCTTTGGGCCATCGTGAACCGCTTAGACCTAGCTGCT 150

QY 42 SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyrAspThrLeu 60
DB 151 TCCGGTATAGACCAACACAGGTCACATCTGACATCTTACATGTGAACAATATTCTGAG 210

QY 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
DB 211 CCATTGATGACGCTAATGACTTCATTCGCCAATGATGAGAGGTGTACTCGTGTGTCAT 270

QY 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
DB 271 AGCTTTGGAGGTTTGAAGTTAGCTTAGCCATGAGTAAGTTCCCGATAAAATCTCTGTC 330

QY 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
DB 331 TCTGTCTTCGTGACTGCAFTATGCGCCGACACAAACACTCACCATCGTTCGTCGAGGAA 390

QY 121 GlyTyrAsnGluArgThrProAlaGluAenTrpLeuAspThrGlnPheLeuProTyrGly 140
DB 391 AGTTTGCAGCAGCATGACACAGAGATGGATGGCTCTGAGCTCGAGACATATGTT 450

QY 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
DB 451 TCAGATAAATTCGGGCTTG---TCTGTGTCTTTCAGCACCAGCACTTCATGAAGCACCCTCTC 507

QY 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
DB 508 TACCACATCTTCCTGTGAGGATCTTGAGCTTGATGCTTCTAAAGAGGCTAGTTCA 567

QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
DB 568 TTGTTTATTAATGAATATATCGAAGATGGAGAACTTTCTGAGAAAGGGTATGATCTGTT 627

QY 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220
DB 628 CCTCGAGCTTACATTGTGTGCAAGAGGACAACATTTATCTCGAAGACCATCAACGATGG 687

QY 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240
DB 688 ATGATCCATAATTTATCCGCGCAATTTAGTGTGATGATGGAAGAGACTCATATATGCCA 747

QY 241 MetLeuCysGluProGlnLysLysCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
DB 748 ATGTTTGGCAACCTCACTACTACTAGTGACCATCTATTGGCAATCGCTGACAAATTC 804

RESULT 8
ADA26531
ID ADA26531 standard; DNA; 950 BP.
XX
AC ADA26531;
XX
XX 20-NOV-2003 (first entry)
XX
DE Arabidopsis thaliana AtSB2L2 gene homologous to tobacco SABP2 gene.
DE
KW salicylic acid-binding protein; SABP2; disease resistance; plant;
KW modulator; tobacco; ds; gene.
KW
```

```
XX Arabidopsis thaliana.
OS
XX WO2003016551-A2.
PN
XX 27-FEB-2003.
PD
XX
XX 16-AUG-2002; 2002WO-US026312.
PF
XX 16-AUG-2001; 2001US-0312863P.
PR
XX (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
PA
XX
XX Klessig DF, Kumar D;
PI
XX WPI; 2003-278578/27.
DR
XX
XX New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for
PT producing transgenic plants with increased resistance to disease, or for
PT screening SABP2 modulators that confer enhanced resistance of plants to
PT disease.
XX
XX Claim 65; Page; 98pp; English.
XX
XX The invention relates to a novel isolated 1079 bp salicylic acid-binding
CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid
CC protein, sequences that specifically hybridize with the nucleic acid, the
CC complement of the nucleic acid or a natural allelic variant of the
CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for
CC conferring increased resistance to disease in plants, or for screening
CC modulators of SABP2, which confer increased or enhanced resistance of
CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also
CC useful for producing transgenic plants with increased resistance to
CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful
CC for studying the molecular mechanisms responsible for the modulation of
CC salicylic acid-mediated disease resistance in plants. This sequence
CC represents the gene encoding the Arabidopsis thaliana SABP2 AtSB2L2 which
CC has homology to the gene (ADA26498) encoding the tobacco SABP2 protein.
CC (Note: this sequence is not given in the specification but is derived
CC from the Genbank accession number given by the inventors).
XX
SQ Sequence 950 BP; 272 A; 185 C; 218 G; 275 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7,54e-71 Length: 950
Score: 695.00 Matches: 138
Percent Similarity: 67.18% Conservative: 36
Best Local Similarity: 53.28% Mismatches: 83
Query Match: 50.80% Indels: 2
DB: 8 Gaps: 2

US-10-780-002-2 (1-260) x ADA26531 (1-950)

QY 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyr 21
DB 45 AAGAGGAGCAACACTTCGTGTAGTACATGCTGCGGCCAGCGCATGGTGGTGTAC 104

QY 22 LysLeuLysProLeuLeuGluAlaGlyHisLysValThrAlaLeuAspLeuAla 41
DB 105 AAGGTTAAGCCCTCTTCGAGGCTTTGGGCCATCGTGTAAACCGCTTAGACCTAGCTGCT 164

QY 42 SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyrAspThrLeu 60
DB 165 TCCGGTATAGACCAACACAGGTCACATCTGACATCTTCTACATGTGAACAATATTCTGAG 224

QY 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
DB 225 CCATTGATGACGCTAATGACTTCATTCGCCAATGATGAGAGGTGTACTCGTGTGTCAT 284

QY 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
DB 285 AGCTTTGGAGGTTTGAAGTTAGCTTAGCCATGAGTAAGTTTCCCGATAAAATCTCTGTC 344
```





QY 244 GluProGlnLysLeuCysAlaSerLeuLeuGluLeuAlaHisLysTyr 259  
 DB 739 AAGCCTCAACACTATGTCTTCTTCTTGGAGATTGCAATATAT 786

## RESULT 10

AB212412  
 ID AB212412 standard; DNA; 792 BP.

AC AB212412;

XX 21-JAN-2003 (first entry)

DE Arabidopsis thaliana stress regulated gene SEQ ID NO 217.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

OS Arabidopsis thaliana.

XX WO200216655-A2.

XX 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US026685.

XX 24-AUG-2000; 2000US-0227866P.

XX 26-JAN-2001; 2001US-02664547P.

XX 22-JUN-2001; 2001US-0300111P.

XX (SCRI ) SCRIPPS RES INST.

XX (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX Harper JF, Kreps J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

PT Identifying a stress condition to which a plant cell has been exposed and  
 PT producing plants with increased tolerance to these abiotic stresses.

PS Claim 144; SEQ ID NO 217; 577pp + Sequence Listing; English.

CC The invention relates to identifying a stress condition to which a plant  
 CC cell has been exposed, comprising: (a) contacting nucleic acid  
 CC representative of expressed polynucleotides in the plant cell with an  
 CC array or probes representative of the plant cell genome; and (b)  
 CC detecting a profile of expressed polynucleotides in the plant cell  
 CC characteristic of a stress response. The method is useful in the  
 CC production of transgenic plants, cells and seeds and in producing plants  
 CC with increased tolerance to abiotic stress. The present sequence is that  
 CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used  
 CC in methods of the invention. Note: The sequence data for this patent is  
 CC not represented in the printed specification but is based on sequence  
 CC information supplied to Derwent by the European Patent Office

XX Sequence 792 BP; 230 A; 167 C; 192 G; 203 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 2,22e-70 Length: 792  
 Score: 690.00 Matches: 134  
 Percent Similarity: 69.92% Conservative: 45  
 Best Local Similarity: 52.34% Mismatches: 75  
 Query Match: 50.44% Indels: 2  
 DB: 6 Gaps: 2

US-10-780-002-2 (1-260) x AB212412 (1-792)

QY 5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTyrLysLeuLys 24

DB 22 CAACACGTCTGTCTAGTACATGGTGTTCCTTCCATGGCGCTGTGTGGTACAGGTAG 81

QY 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaSerGlyThr 44

DB 82 CCGCAGCTCGAGGCTTCTGGCCACCGCGTAACCGCGGTAGATCTAGTCCCTCCGGTATA 141

QY 45 AspLeu---ArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMet 63  
 DB 142 GACATGACCAGGTCAATCACAGATATATCCACATCGGAACATATCTCAGAGCCATTGATG 201  
 QY 64 GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 83  
 DB 202 CAGCTTAATGACCTCACTACCGATGATGAGAAGGTTGTCTTGTGGTCATAGCTAGGA 261  
 QY 84 GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPhe 103  
 DB 262 GGTITGAGTTAGCTATGGCCATGATATGTTCCGACCAAAATCTCTGTTCTGTCTTT 321  
 QY 104 LeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsn 123  
 DB 322 GTGACTGTCTATGATCCAGACACCAACACTCCACCATCTTCGTATGGATAAGCTAAGA 381  
 QY 124 GluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGlu 143  
 DB 382 AAAGAAACTTTCACGAGAGGAATGGTTAGACACCGTGTATTACGAGC---GAGAAACCTGAT 438  
 QY 144 GluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeu 163  
 DB 439 TTTCTTAGCGAGTTTGGATTTTTCGACCAAGAAATTCATGGCCCAAGAACTTGTATCAGTTG 498  
 QY 164 CysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMet 183  
 DB 499 TCTCAGTCCAGATCTTGAATTCGCGAAATGTTGGTGGGCGCAACCCATTGATTAG 558  
 QY 184 GluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgVal 203  
 DB 559 AAAGATATGGCAGAGAGAAGAGCTTCAGTAGGAAGGATACGGATCCGTTACACGTATA 618  
 QY 204 TyrIleValCysThrGluAspLysGlyIleProGluGluPheGluArgTrpGlnIleAsp 223  
 DB 619 TTTATTGTATGCGGAAGGATCTTGTGTACCCCGAAGATTACGAGCGATCGATGATCAGC 678  
 QY 224 AsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243  
 DB 679 AACTTTCCCAAAAGAGTAATGAGATCAAGACGCGAGATCATATGCCAATGTTCTCC 738  
 QY 244 GluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259  
 DB 739 AAGCCTCAACACTATGTGTCTTCTCTTGGAGATTGCAATATAT 786  
 RESULT 11  
 ADN74620  
 ID ADN74620 standard; cDNA; 792 BP.  
 XX AC ADN74620;  
 XX DT 15-JUL-2004 (first entry)  
 XX DE Thale cress cDNA repressed in E2Fa/Dpa expressing plants SeqID 2515.  
 XX KW gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;  
 XX KW growth regulator; animal feed product; thale cress;  
 XX KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.  
 XX OS Arabidopsis thaliana.  
 XX PN WO2004035798-A2.  
 XX PD 29-APR-2004.  
 XX PF 20-OCT-2003; 2003WO-EP011658.  
 XX PR 18-OCT-2002; 2002EP-00079408.  
 XX PA (CROP-) CROPDESIGN NV.  
 XX PI Inze D, De Veylder L, Vlieghe K;  
 XX

WPI; 2004-348456/32.  
P-PSDB; ADN74621.

Altering plant characteristics, useful for producing plants for enzyme or pharmaceutical production comprises modifying in a plant, expression of one or more nucleic acids and/or modifying level or activity of one or more proteins.

Claim 1; SEQ ID NO 2515; 134pp; English.

This invention relates to a novel method for altering one or more plant characteristics. Specifically, it refers to identifying genes that are up- or down-regulated in transgenic plants overexpressing the heterodimeric E2Fa/DPA transcription factor of Arabidopsis and using these sequences to alter plant characteristics accordingly. The present invention describes generating transgenic plants for the production of growth regulators, enzymes, therapeutics, pharmaceuticals and animal feed products, where the altered plant characteristics are selected from increased yield or biomass, enhanced survival capacity, stress tolerance, plant architecture or physiology, altered endoreplication, biochemistry, signal transduction, storage lipid mobilisation and/or altered photosynthesis, each relative to the corresponding wild type plants. Accordingly, these sequences can also be useful as positive or negative selectable markers during transformation of cells or tissues. The identified genes play a role in a variety of biological processes such as DNA replication, cell wall biosynthesis, nitrogen and/or carbon metabolism or they function as transcription factors. This polynucleotide sequence is thale cress CDNA repressed i.3 fold or more in plants overexpressing the E2Fa/DPA transcription factor, given in an exemplification of the invention.

Sequence 792 BP; 230 A; 167 C; 192 G; 203 T; 0 U; 0 Other;

| Alignment Scores:      |          |               |     |
|------------------------|----------|---------------|-----|
| Pred. No.:             | 2,228-70 | Length:       | 792 |
| Score:                 | 690.00   | Matches:      | 134 |
| Percent Similarity:    | 69.92%   | Conservative: | 75  |
| Best Local Similarity: | 52.34%   | Mismatches:   | 45  |
| Query Match:           | 50.44%   | Indels:       | 2   |
| DB:                    | 12       | Gaps:         | 2   |

US-10-780-002-2 (1-260) x ADN74620 (1-792)

[illegible]

|    |     |   |     |
|----|-----|---|-----|
| Qy | 164 | CysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMet    | 183 |
|    |     | :::   |     |
| Db | 499 | TCTCCAGTCCAAGATCTTGAATTTGGCGAAAAATGTTGGTCAGCGCAACCCATTTGATTAAAG | 558 |
|    |     | :::   |     |
| Qy | 184 | GluAspLeuSerLysAlaLysTyPheThrAspGluArgPheGlySerValLysArgVal     | 203 |
|    |     | :::   |     |
| Db | 559 | AAAGATATGCGACAGACAGAGACTTTCAGTCAGGAGGATACGGATCCGTTACACGTATA     | 618 |
|    |     | :::   |     |
| Qy | 204 | TyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTyrGlnIleAsp    | 223 |
|    |     | :::   |     |
| Db | 619 | TTTATTGTATGTCGGAAGAGATCTTGTGTACCCCAAGATTACCAGCGATTCGATGATCAGC   | 678 |
|    |     | :::   |     |
| Qy | 224 | AsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys    | 243 |
|    |     | :::   |     |
| Db | 679 | AACTTTCCCCAAAGAAGATGTAATGGAGATCAAGACGCAGATCATATGCCAATGTTCTTCC   | 738 |
|    |     | :::   |     |
| Qy | 244 | GluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr                | 259 |
|    |     | :::   |     |
| Db | 719 | AAGCTCTCAACAACATATGTGCTCTCTCTTGAGATTCGAATAAATAAT                | 786 |
|    |     | :::   |     |

RESULT 12

|          |   |
|----------|---|
| ADA26532 |   |
| ID       | ADA26532 standard; DNA; 1152 BP.  |
| XX       |   |
| XX       | ADA26532;   |
| XX       |   |
| DT       | 20-NOV-2003 (first entry)   |
| XX       |   |
| DE       | Arabidopsis thaliana A15B2L3 gene homologous to tobacco SABP2 gene.       |
| XX       |   |
| KW       | salicylic acid-binding protein; SABP2; disease resistance; plant;         |
| KW       | modulator; tobacco; ds; gene.   |
| XX       |   |
| OS       | Arabidopsis thaliana.   |
| XX       |   |
| FN       | W02003016551-A2.  |
| XX       |   |
| PD       | 27-FEB-2003.  |
| XX       |   |
| PF       | 16-AUG-2002; 2002WO-US026312.   |
| XX       |   |
| PR       | 16-AUG-2001; 2001US-0312863P.   |
| XX       |   |
| PA       | (BOYC-) BOYCE THOMPSON INST PLANT RES INC.                                |
| XX       |   |
| PI       | Klessig DF, Kumar D;  |
| XX       |   |
| DR       | WPI; 2003-278578/27.  |
| XX       |   |
| PT       | New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for |
| PT       | producing transgenic plants with increased resistance to disease, or for  |
| PT       | screening SABP2 modulators that confer enhanced resistance of plants to   |
| PT       | disease.  |
| XX       |   |
| PS       | Claim 65; Page; 98pp; English.  |
| XX       |   |
| CC       | The invention relates to a novel isolated 1079 bp salicylic acid-binding  |
| CC       | protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid         |
| CC       | protein, sequences that specifically hybridize with the nucleic acid, the |
| CC       | complement of the nucleic acid or a natural allelic variant of the        |
| CC       | sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for      |
| CC       | conferring increased resistance to disease in plants, or for screening    |
| CC       | modulators of SABP2, which confer increased or enhanced resistance of     |
| CC       | plants to disease. The SABP2 or SABP2 homologue nucleic acid is also      |
| CC       | useful for producing transgenic plants with increased resistance to       |
| CC       | disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful  |
| CC       | for studying the molecular mechanisms responsible for the modulation of   |
| CC       | salicylic acid-mediated disease resistance in plants. This sequence       |
| CC       | represents the gene encoding the Arabidopsis thaliana SABP2 A15B2L3 which |
| CC       | has homology to the gene (ADA26498) encoding the tobacco SABP2 protein.   |
| CC       | (Note: this sequence is not given in the specification but is derived     |
| CC       | from the Genbank accession number given by the inventors).                |

```
XX SQ Sequence 1152 BP; 372 A; 214 C; 253 G; 313 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3.8e-70 Length: 1152
Score: 690.00 Matches: 134
Percent Similarity: 69.92% Conservativity: 45
Best Local Similarity: 52.34% Mismatches: 75
Query Match: 50.44% Indels: 2
DB: 8 Gaps: 2

US-10-780-002-2 (1-260) x ADA26532 (1-1152)
QY 5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTyrTyrLeuLeuLys 24
DB 151 CAACAGCTGGTTCATAGTACATGGTGTCCATGGCGCTGTGTGGTACAGGTAG 210
QY 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaSerGlyThr 44
DB 211 CCGCAGCTCGAGGCTTCTGCCACCGCGTAACCGCGTAGATCTAGCTCGGTATA 270
QY 45 AspLeu---ArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMet 63
DB 271 GACATGACGAGGTCATACAGATATATCCATGCGAACAATACTCAGAGCCATTGATG 330
QY 64 GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 83
DB 331 CAGCTAATGACCTCATACCATGATGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGA 390
QY 84 GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaValPhe 103
DB 391 GGTTCAGTTAGTATGCTATGCTATGATATGTTTCCGACCAAAATCTCTGTTCTGTT 450
QY 104 LeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsn 123
DB 451 GTGACTGCTATGATGCGACACACCAACACTCACCATCTCTGATGGGATAAGCTAAGA 510
QY 124 GluArgThrProAlaGlnAsnTyrLeuAspThrGlnPheLeuProTyrGlySerProGlu 143
DB 511 AAAGAACTTCAGGAGAGAGAGTGTAGACCGTGTGTGTGTGTGTGTGTGTGTGTGT 567
QY 144 GluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeu 163
DB 568 TTTCTCTAGCGAGTTTGGATTTTGGACCAAGAATTCATGCCCAAGAACTTGTATCAGTTG 627
QY 164 CysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMet 183
DB 628 TCTCCAGTCCAGAGATCTGAATTTGGCGAAAATGTTGTGAGGCGCAAAACCATTTAAG 687
QY 184 GluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgVal 203
DB 688 AAAGATATGCGAGAGAGAGAGCTTCAGTGGAGAGGATACGGATCCGTTACAGGTATA 747
QY 204 TyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTyrGlnIleAsp 223
DB 748 TTTATTGTATGCGAAGAGTCTGTGTCAACCGAAGATTACCAGCATCGATGATCAGC 807
QY 224 AsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243
DB 808 AACTTTTCCCAAAAGAGTAATGGAGATCAAAAGACGACAGATCATATGCGCAATGTTCTCC 867
QY 244 GluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
DB 868 AAGCCTCAAACTATGTGCTCTTCTCTTGGAGATTGCAATATAATAT 915

RESULT 13
ID ADA26533
XX ADA26533 standard; DNA; 1109 BP.
XX AC ADA26533;
XX DT 20-NOV-2003 (first entry)
XX
```

DE Arabidopsis thaliana AtSB2L4 gene homologous to tobacco SABP2 gene.

XX salicylic acid-binding protein; SABP2; disease resistance; plant;

KW modulator; tobacco; ds; gene.

XX Arabidopsis thaliana.

OS WO2003016551-A2.

PN 27-FEB-2003.

PD 16-AUG-2002; 2002WO-US026312.

PF 16-AUG-2001; 2001US-0312863P.

XX (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

PA Klessig DF, Kumar D;

PI WPI; 2003-278578/27.

DR New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for

PT producing transgenic plants with increased resistance to disease, or for

PT screening SABP2 modulators that confer enhanced resistance of plants to

PT disease.

XX Claim 65; Page; 98pp; English.

XX The invention relates to a novel isolated 1079 bp salicylic acid-binding

CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid

CC protein, sequences that specifically hybridize with the nucleic acid, the

CC complement of the nucleic acid or a natural allelic variant of the

CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for

CC conferring increased resistance to disease in plants, or for screening

CC modulators of SABP2, which confer increased or enhanced resistance of

CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also

CC useful for producing transgenic plants with increased resistance to

CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful

CC for studying the molecular mechanisms responsible for the modulation of

CC salicylic acid-mediated disease resistance in plants. This sequence

CC represents the gene encoding the Arabidopsis thaliana SABP2 AtSB2L4 which

CC has homology to the gene (ADA26498) encoding the tobacco SABP2 protein.

CC (Note: this sequence is not given in the specification but is derived

CC from the Genbank accession number given by the inventors).

XX SQ Sequence 1109 BP; 365 A; 193 C; 235 G; 316 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.61e-69 Length: 1109

Score: 680.50 Matches: 134

Percent Similarity: 67.58% Conservativity: 39

Best Local Similarity: 52.34% Mismatches: 82

Query Match: 49.74% Indels: 1

DB: 8 Gaps: 1

US-10-780-002-2 (1-260) x ADA26533 (1-1109)

QY 5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTyrTyrLeuLys 24

DB 19 AAGCGGTTTGTCTCGTCCATGGGTGTGCCACGCGCTGTGTGTGTGTGTGTGTGTGTG 78

QY 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaSerGlyThr 44

DB 79 ACGCATCTGGAGGCTGTAGTCTGTGTGACCGCGGTGATCTAGCTGCAATCCGGTATA 138

QY 45 AspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMetGlu 64

DB 139 AATATGACTAGATTGGAAGAGATTTCAGACTTTGAAGGATTACTGCAACCTTTGCTTGA 198

QY 65 LeuMetGluSerLeu---SerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 83

DB 199 TTACTGAACCTCACTTGGCTCGGATGACGATGAAGGTGATCTTGTGTGCGCATAGTAGGA 258



QY 84 GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysValleTyrAlaAlaValPhe 103  
 DB 259 GGAATACCTGCTGCTCTGACATATTCCTAGTAAGATTGCTACTATTGTTTC 318  
 QY 104 LeuAlaAlaPheMetProAspSerValHisAsnSerPheValLeuGluGlnTyrAsn 123  
 DB 319 TTGACAGCTTTTATGCCGACACAGAACCTACCTGCTTATGTTTACCAAAAGCTAATC 378  
 QY 124 GluA-gThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGlu 143  
 DB 379 AGAAGCGTTTCCACAAGAGGATGTTGGACACCGTGTGGAAACGTATGGGAACATGAA 438  
 QY 144 GluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeu 163  
 DB 439 TGTCTCTAGAGTTTGCTCTTTTGGACCAAAAGTTCATGGCCCAAGAAATTTGTATCACTC 498  
 QY 164 CysSerProGluAspLeuAlaLaserSerLeuValArgProSerSerLeuPheMet 183  
 DB 499 TCTCCGTCACAGATCTTGAATGGGAAAATGTTGGTGAAGATTAACCCCATATTACA 558  
 QY 184 GluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgVal 203  
 DB 559 AATAATCTGGAGGACAAAGAGCTTTAGTGAGGAAGGTACCGTTACCGTTACAGCTATA 618  
 QY 204 TyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAsp 223  
 DB 619 TATATTGTATGTGGAGAGACATGCGGTACCCGAGGATTACCATGGTGGATGATCAAG 678  
 QY 224 AsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243  
 DB 679 AACTTCCCAAAAGAGTATGAGATCAATATGTCAGATCATATGGCAATGTTCTCC 738  
 QY 244 GluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259  
 DB 739 AAGCCTCAAACTATGTCTCTCTCGTGGAGATTGCGATGTAATAT 786  
 RESULT 14  
 ADA26535  
 ID ADA26535 standard; DNA; 1046 BP.  
 XX  
 AC ADA26535;  
 XX  
 DT 20-NOV-2003 (first entry)  
 XX  
 DE Arabidopsis thaliana AtSB2L6 gene homologous to tobacco SABP2 gene.  
 XX  
 KW salicylic acid-binding protein; SABP2; disease resistance; plant;  
 KW modulator; tobacco; ds; gene.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN W02003016551-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 PF 16-AUG-2002; 2002WO-US026312.  
 XX  
 PR 16-AUG-2001; 2001US-0312863P.  
 XX  
 PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.  
 XX  
 PI Klessig DF, Kumar D;  
 XX  
 DR WPI; 2003-278578/27.  
 XX  
 PT New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for  
 PT producing transgenic plants with increased resistance to disease, or for  
 PT screening SABP2 modulators that confer enhanced resistance of plants to  
 PT disease.  
 XX  
 PS Claim 65; Page; 98pp; English.  
 XX  
 CC The invention relates to a novel isolated 1079 bp salicylic acid-binding

CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid  
 CC protein, sequences that specifically hybridize with the nucleic acid, the  
 CC complement of the nucleic acid or a natural allelic variant of the  
 CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for  
 CC conferring increased resistance to disease in plants, or for screening  
 CC modulators of SABP2, which confer increased or enhanced resistance of  
 CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also  
 CC useful for producing transgenic plants with increased resistance to  
 CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful  
 CC for studying the molecular mechanisms responsible for the modulation of  
 CC salicylic acid-mediated disease resistance in plants. This sequence  
 CC represents the gene encoding the Arabidopsis thaliana SABP2 AtSB2L6 which  
 CC has homology to the gene (ADA26498) encoding the tobacco SABP2 protein.  
 CC (Note: this sequence is not given in the specification but is derived  
 CC from the Genbank accession number given by the inventors).  
 XX

SQ Sequence 1046 BP; 330 A; 195 C; 248 G; 273 T; 0 U; 0 Other;

Alignment Scores: 2.37e-67 Length: 1046  
 Pred. No.: 665.50 Matches: 128  
 Score: 67.57% Conservative: 47  
 Percent Similarity: 67.57% Mismatches: 83  
 Best Local Similarity: 49.42% Indels: 1  
 Query Match: 48.65% Gaps: 1  
 DB: 8

US-10-780-002-2 (1-260) x ADA26535 (1-1046)

QY 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyTyrSerTrpTyr 21  
 DB 31 AAGAACCAAGACGGTTTGTCTCATCCATGGGTGTGCCACGGGGCATGACTTGGGAC 90  
 QY 22 LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla 41  
 DB 91 AAGGTAAACCGCAGCTGGAGTTGCGAGTCACTGTGTGACGGCATGGATCTTGTCTCA 150  
 QY 42 SerGlyThrAspLeuArgLysIleGluLeuArgThrLeuTyrAspTyrThrLeuPro 61  
 DB 151 TCAGTATAACATGACCAAGTGGAGAGATTTCAGACTCTGAACGATTACTGCNAACCA 210  
 QY 62 LeuMetGluLeuMetGluSerLeuSerAlaAspGlu---LysValIleLeuValGlyHis 80  
 DB 211 TTGCTTGAGTTTCTGAGCTCGCTTGGCTCAGATGACGGTAAGTGATTGTTGTGCTCAT 270  
 QY 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100  
 DB 271 AGCATGGGAGGAATATCCGCTGCACTTCTCTCAGACGCTTCGCTTGAAGATTGCCGCT 330  
 QY 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120  
 DB 331 ATTGTCTTTTGGACAGCTTTCATGCCGACACAATAAACCCACCCTGCTTATGTTTACGAA 390  
 QY 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140  
 DB 391 AAGCTGCTCAGACGATTCACAGAGGAATGTTGGACACACCGCTGTGTGAATACGGG 450  
 QY 141 SerProGluGluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeu 160  
 DB 451 AAACCTGATTTTCTCTCTACAGTATATCTTTTGGGACCAAAAGTTTATCGCCAAAGAAATG 510  
 QY 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAspSerLeuValArgProSerSer 180  
 DB 511 TATCAAAACTCTCCAGTTCAAGATCTTGAAGTGGTGAAGACATATTAGTAGGGGAAACCCG 570  
 QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200  
 DB 571 TTAGTTACAAACAATCTGGCAGGACAAAGAGCTTTAGTAGGAAGGCTACGATCCGTT 630  
 QY 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220  
 DB 631 ACACGTATATATATTGTATGCAGAGGAGTCTTGTGGAAAGTCGAAGATTACACGGTTGG 690  
 QY 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240



|    |     |  |     |    |              |                |
|----|-----|--|-----|----|--------------|----------------|
| Db | 691 | ATGATTAGCAACTTTCACCAAGAAAGTAATGGAGATCAAATGTGCAGATCATATGCCA | 750 | PR | 16-JUN-1999; | 99US-0139453P. |
|    |     | :::  |     | PR | 17-JUN-1999; | 99US-0139492P. |
|    |     | :::  |     | PR | 18-JUN-1999; | 99US-0139454P. |
| Qy | 241 | MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluLeuHisLysTyr     | 259 | PR | 18-JUN-1999; | 99US-0139455P. |
|    |     | :::  |     | PR | 18-JUN-1999; | 99US-0139456P. |
| Db | 751 | ATGTTCTCCAAAGCCTCAAGAAGTTTGTCCTTCTCTGGAGATTGCAATAAATAT     | 807 | PR | 18-JUN-1999; | 99US-0139457P. |
|    |     | :::  |     | PR | 18-JUN-1999; | 99US-0139458P. |
|    |     | :::  |     | PR | 18-JUN-1999; | 99US-0139459P. |
|    |     | :::  |     | PR | 18-JUN-1999; | 99US-0139460P. |
|    |     | :::  |     | PR | 18-JUN-1999; | 99US-0139461P. |
|    |     | :::  |     | PR | 18-JUN-1999; | 99US-0139462P. |
|    |     | :::  |     | PR | 18-JUN-1999; | 99US-0139463P. |
|    |     | :::  |     | PR | 18-JUN-1999; | 99US-0139750P. |
|    |     | :::  |     | PR | 18-JUN-1999; | 99US-0139763P. |
|    |     | :::  |     | PR | 21-JUN-1999; | 99US-0139817P. |
|    |     | :::  |     | PR | 22-JUN-1999; | 99US-0139899P. |
|    |     | :::  |     | PR | 23-JUN-1999; | 99US-0140353P. |
|    |     | :::  |     | PR | 23-JUN-1999; | 99US-0140354P. |
|    |     | :::  |     | PR | 24-JUN-1999; | 99US-0140695P. |
|    |     | :::  |     | PR | 28-JUN-1999; | 99US-0140823P. |
|    |     | :::  |     | PR | 29-JUN-1999; | 99US-0140991P. |
|    |     | :::  |     | PR | 30-JUN-1999; | 99US-0141287P. |
|    |     | :::  |     | PR | 01-JUL-1999; | 99US-0141842P. |
|    |     | :::  |     | PR | 01-JUL-1999; | 99US-0142154P. |
|    |     | :::  |     | PR | 02-JUL-1999; | 99US-0142055P. |
|    |     | :::  |     | PR | 06-JUL-1999; | 99US-0142390P. |
|    |     | :::  |     | PR | 08-JUL-1999; | 99US-0142803P. |
|    |     | :::  |     | PR | 09-JUL-1999; | 99US-0142920P. |
|    |     | :::  |     | PR | 12-JUL-1999; | 99US-0142977P. |
|    |     | :::  |     | PR | 13-JUL-1999; | 99US-0143542P. |
|    |     | :::  |     | PR | 14-JUL-1999; | 99US-0143624P. |
|    |     | :::  |     | PR | 15-JUL-1999; | 99US-0144005P. |
|    |     | :::  |     | PR | 16-JUL-1999; | 99US-0144085P. |
|    |     | :::  |     | PR | 16-JUL-1999; | 99US-0144086P. |
|    |     | :::  |     | PR | 19-JUL-1999; | 99US-0144325P. |
|    |     | :::  |     | PR | 19-JUL-1999; | 99US-0144331P. |
|    |     | :::  |     | PR | 19-JUL-1999; | 99US-0144332P. |
|    |     | :::  |     | PR | 19-JUL-1999; | 99US-0144333P. |
|    |     | :::  |     | PR | 19-JUL-1999; | 99US-0144334P. |
|    |     | :::  |     | PR | 19-JUL-1999; | 99US-0144335P. |
|    |     | :::  |     | PR | 20-JUL-1999; | 99US-0144352P. |
|    |     | :::  |     | PR | 20-JUL-1999; | 99US-0144632P. |
|    |     | :::  |     | PR | 20-JUL-1999; | 99US-0144884P. |
|    |     | :::  |     | PR | 21-JUL-1999; | 99US-0144814P. |
|    |     | :::  |     | PR | 21-JUL-1999; | 99US-0145086P. |
|    |     | :::  |     | PR | 21-JUL-1999; | 99US-0145088P. |
|    |     | :::  |     | PR | 22-JUL-1999; | 99US-0145085P. |
|    |     | :::  |     | PR | 22-JUL-1999; | 99US-0145087P. |
|    |     | :::  |     | PR | 22-JUL-1999; | 99US-0145089P. |
|    |     | :::  |     | PR | 22-JUL-1999; | 99US-0145192P. |
|    |     | :::  |     | PR | 23-JUL-1999; | 99US-0145145P. |
|    |     | :::  |     | PR | 23-JUL-1999; | 99US-0145218P. |
|    |     | :::  |     | PR | 23-JUL-1999; | 99US-0145224P. |
|    |     | :::  |     | PR | 26-JUL-1999; | 99US-0145276P. |
|    |     | :::  |     | PR | 27-JUL-1999; | 99US-0145913P. |
|    |     | :::  |     | PR | 27-JUL-1999; | 99US-0145918P. |
|    |     | :::  |     | PR | 27-JUL-1999; | 99US-0145919P. |
|    |     | :::  |     | PR | 28-JUL-1999; | 99US-0145951P. |
|    |     | :::  |     | PR | 02-AUG-1999; | 99US-0146386P. |
|    |     | :::  |     | PR | 02-AUG-1999; | 99US-0146388P. |
|    |     | :::  |     | PR | 02-AUG-1999; | 99US-0146389P. |
|    |     | :::  |     | PR | 03-AUG-1999; | 99US-0147038P. |
|    |     | :::  |     | PR | 04-AUG-1999; | 99US-0147204P. |
|    |     | :::  |     | PR | 04-AUG-1999; | 99US-0147302P. |
|    |     | :::  |     | PR | 05-AUG-1999; | 99US-0147192P. |
|    |     | :::  |     | PR | 05-AUG-1999; | 99US-0147260P. |
|    |     | :::  |     | PR | 06-AUG-1999; | 99US-0147303P. |
|    |     | :::  |     | PR | 06-AUG-1999; | 99US-0147416P. |
|    |     | :::  |     | PR | 09-AUG-1999; | 99US-0147493P. |
|    |     | :::  |     | PR | 09-AUG-1999; | 99US-0147935P. |
|    |     | :::  |     | PR | 10-AUG-1999; | 99US-0148171P. |
|    |     | :::  |     | PR | 11-AUG-1999; | 99US-0148319P. |
|    |     | :::  |     | PR | 12-AUG-1999; | 99US-0148341P. |

RESULT 15

AAC49239  
ID AAC49239 standard; DNA; 984 BP.  
XX AAC49239;  
AC  
XX  
DT 18-OCT-2000 (first entry)  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 60437.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 12, 2005, 17:29:58 ; Search time 3425 Seconds  
(without alignments)  
3678.352 Million cell updates/sec

Title: US-10-780-002-2

Perfect score: 1368

Sequence: 1 MKEGRHFVVLHGACGGGWSW.....MLCEPQKLCASLLEIAHKYN 260

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xl  
-Q=/cgn2\_1/USFTO.spool\_p/US10780002/runat\_11092005\_104747\_26597/app.query.fasta\_1.455  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.0.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10780002\_@CNG\_1\_1\_5600 @runat\_11092005\_104747\_26597 -NCFU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID         | Description |
|------------|-------|-------------|--------|------------|-------------|
| 1          | 1368  | 100.0       | 1079   | 8 AY485932 | Nicotiana   |
| 2          | 811   | 59.3        | 961    | 8 AF178576 | Rauwolfia   |
| 3          | 793   | 58.0        | 1077   | 8 AF269158 | Citrus si   |
| 4          | 768   | 56.1        | 792    | 6 AX506625 | Sequence    |

| 5  | 768   | 56.1 | 792    | 8 | BT014881  | Arabidops  |
|----|-------|------|--------|---|-----------|------------|
| 6  | 767   | 56.1 | 935    | 8 | BT002859  | Arabidops  |
| 7  | 734   | 53.7 | 891    | 8 | AY751530  | Catharant  |
| 8  | 696   | 50.9 | 870    | 8 | AY086590  | Arabidops  |
| 9  | 695   | 50.8 | 792    | 6 | AX506856  | Sequence   |
| 10 | 695   | 50.8 | 792    | 6 | AY056068  | Arabidops  |
| 11 | 695   | 50.8 | 932    | 8 | AF361627  | Arabidops  |
| 12 | 690   | 50.4 | 792    | 6 | CQ806104  | Sequence   |
| 13 | 690   | 50.4 | 792    | 6 | AX505522  | Sequence   |
| 14 | 690   | 50.4 | 792    | 8 | BT015031  | Arabidops  |
| 15 | 690   | 50.4 | 878    | 8 | BT014901  | Arabidops  |
| 16 | 680.5 | 49.7 | 792    | 8 | BT010429  | Arabidops  |
| 17 | 680.5 | 49.7 | 902    | 8 | AK175772  | Arabidops  |
| 18 | 669.5 | 48.9 | 1073   | 8 | AK061058  | Oryza sat  |
| 19 | 661.5 | 48.4 | 924    | 8 | BT004125  | Arabidops  |
| 20 | 645   | 47.1 | 783    | 8 | BT012414  | Arabidops  |
| 21 | 645   | 47.1 | 944    | 8 | BT010709  | Arabidops  |
| 22 | 640.5 | 46.8 | 770    | 6 | AX51280   | Sequence   |
| 23 | 640.5 | 46.8 | 771    | 6 | AX506583  | Sequence   |
| 24 | 640.5 | 46.8 | 771    | 8 | BT006227  | Arabidops  |
| 25 | 640.5 | 46.8 | 928    | 8 | AK117107  | Arabidops  |
| 26 | 635.5 | 46.5 | 123360 | 8 | AC003040  | Arabidops  |
| 27 | 635   | 46.4 | 777    | 8 | AY093714  | Arabidops  |
| 28 | 635   | 46.4 | 808    | 8 | AY142490  | Arabidops  |
| 29 | 634   | 46.3 | 928    | 8 | AY046015  | Arabidops  |
| 30 | 632.5 | 46.2 | 795    | 6 | AX654444  | Sequence   |
| 31 | 629   | 46.0 | 911    | 8 | AY058115  | Arabidops  |
| 32 | 605   | 44.2 | 789    | 8 | AY455313  | Lycopersi  |
| 33 | 590   | 43.1 | 828    | 8 | AY142031  | Arabidops  |
| 34 | 590   | 43.1 | 904    | 8 | AY074858  | Arabidops  |
| 35 | 589   | 43.1 | 903    | 8 | AY086981  | Arabidops  |
| 36 | 585   | 42.8 | 1078   | 8 | HB040402  | Hevea bras |
| 37 | 585   | 42.8 | 1091   | 6 | AS9586    | Sequence 1 |
| 38 | 574.5 | 42.0 | 978    | 6 | AX653408  | Sequence   |
| 39 | 571.5 | 41.8 | 198780 | 8 | ATC819V86 | Arabidops  |
| 40 | 571.5 | 41.8 | 206420 | 8 | ATAP21    | Arabidops  |
| 41 | 564   | 41.2 | 57180  | 8 | ATF18D22  | Arabidops  |
| 42 | 554.5 | 40.5 | 1087   | 8 | BT012867  | Lycopersi  |
| 43 | 548   | 40.1 | 860    | 6 | AX654758  | Sequence   |
| 44 | 548   | 40.1 | 948    | 8 | OSPIR7BR  | Arabidops  |
| 45 | 540.5 | 39.5 | 142239 | 8 | AF003266  | Oryza sat  |

#### ALIGNMENTS

| RESULT 1   | AY485932  | Nicotiana tabacum salicylic acid-binding protein 2 mRNA | linear | PLN 07-JAN-2004 |
|------------|---|---|--------|-----------------|
| LOCUS      | AY485932  | 1079 bp   | mRNA   |                 |
| DEFINITION | Nicotiana tabacum salicylic acid-binding protein 2 mRNA, complete cds.  |   |        |                 |
| ACCESSION  | AY485932  |   |        |                 |
| VERSION    | AY485932.1  | GI:40549302   |        |                 |
| KEYWORDS   | Nicotiana tabacum (common tobacco)  |   |        |                 |
| SOURCE     | Nicotiana tabacum   |   |        |                 |
| ORGANISM   | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Nicotiana. |   |        |                 |
| REFERENCE  | 1 (bases 1 to 1079)   |   |        |                 |
| AUTHORS    | Kumar,D. and Kleissig,D.F.  |   |        |                 |
| TITLE      | High-affinity salicylic acid-binding protein 2 is required for plant innate immunity and has salicylic acid-stimulated lipase activity  |   |        |                 |
| JOURNAL    | Proc. Natl. Acad. Sci. U.S.A. 100 (26), 16101-16106 (2003)  |   |        |                 |
| PUBMED     | 14673096  |   |        |                 |
| REFERENCE  | 2 (bases 1 to 1079)   |   |        |                 |
| AUTHORS    | Kumar,D. and Kleissig,D.F.  |   |        |                 |
| TITLE      | Direct Submission   |   |        |                 |
| JOURNAL    | Submitted (24-NOV-2003) BTI, Tower Rd., Ithaca, NY 14853, USA   |   |        |                 |
| FEATURES   | Location/Qualifiers   |   |        |                 |
| source     | 1..1079   |   |        |                 |
|            | /organism="Nicotiana tabacum"   |   |        |                 |
|            | /mol_type="mRNA"  |   |        |                 |



```
Db 156 GCGGCGCATCAACCAAGAGGCTCGATGAGATTACACACATTTTCGGGACTACTCGGAGCCC 215
Qy 62 LeuMetGluLeuMetGluSerLeuSerAlaAspGluValIleValLeuValGlyHisSer 81
Db 216 TTGATGGAAGTCATGCTAGTATCTCTCTGATGAGAGGTTGTTCTTCTTGCCCATAGC 275
Qy 82 LeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLeuTyProGlnHisIleTyAlaAla 101
Db 276 TTTGGTGGCATGAGTTTGGTCTTGCCATGGAAACCTCCAGAGAAGATATCAGTTGCT 335
Qy 102 ValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGln 121
Db 336 GTTTTATTTGCTGCAATGATGCTGATCTCTTAACACTCACTTATCCGTTTGAGAAG 395
Qy 122 TyrAsnGluArgThrProAlaGluAsnTTPLeuAspThrGlnPheLeuProTyGlySer 141
Db 396 TACATGAGAAGTGTCCGACGATATGATGTTGGACTCACAGTTTCAACCTACGGAAC 455
Qy 142 ProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisIleTyLeu 161
Db 456 CCAGACAACCCAGGAATGTCATGATTCTTGGACCTCAGTTTATGGCCCTCAAAATGTTTC 515
Qy 162 GlnLeuCySerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeu 181
Db 516 CAGAAATTGCTCAGTCAGGACCTTGAATTTAGCCAAATGTTGACTCGACAGGTTGTTA 575
Qy 182 PheMetGluAspLeuSerLysAlaLysTyPheThrAspGluArgPheGlySerValLys 201
Db 576 TTTTCCAGATTGTCAGGAGGCCAAAGATTTCTCAACCGAGAGTACGTTCCGTTGAAG 635
Qy 202 ArgValTyIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgThrGln 221
Db 636 CGAGCTTATATCTTTTGCATGAAGATAAATCATTTCCAGTTGAGTTTCAGAAATGGTTT 695
Qy 222 IleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMet 241
Db 696 GTTGAAGTGTGGAGCTGATTAAGTAAAGAAATCAAAAGAGCAGATCATATGGAATG 755
Qy 242 LeuCyGluProGlnLysLeuCyAlaSerLeuLeuGluIleAla 256
Db 756 CTTTCAGCAAGGGAAGTTTGCAAGTGCCTGCTTGATATATCA 800

AF269158 1077 bp mRNA linear PLN 02-AUG-2002
Citrus sinensis ethylene-induced esterase mRNA, complete cds.
AF269158
AF269158.1 GI:14279436
.
Citrus sinensis
Citrus sinensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Sapindales; Rutaceae; Citrus.
1 (bases 1 to 1077)
Zhong,G.Y., Goren,R., Rivov,J., Sisler,E.C. and Holland,D.
Characterization of an ethylene-induced esterase gene isolated from
Citrus sinensis by competitive hybridization
Physiol. Plantarum 113 (2), 267-274 (2001)
12060305
2 (bases 1 to 1077)
Zhong,G.Y., Goren,R., Rivov,J. and Holland,D.
Direct Submission
Submitted (18-MAY-2000) Horticulture, Faculty of Agricultural, Food
and Environmental Quality Sciences, The Hebrew University of
Jerusalem, Rehovot 76100, Israel
Location/Qualifiers
1. .1077
/organism="Citrus sinensis"
/mol_type="mRNA"
/db_xref="taxon:2711"
55. .858
/notes="EIE"
```

```
/codon_start=1
/product="ethylene-induced esterase"
/protein_id="AAKS8599.1"
/db_xref="GI:14279437"
/translation="MEEVVGMEEKHFVLVHGVNHCANWCYKILKARLVAGGHRVTAVDL
AASGNNMRIEDVHTFHAYSEVMEVLASLPAEEKVILVHSLHSGVTIALLAGDKPEPHK
ISVAVFTAFNPDTHRPSFVLEQYSEKMGKEDDSWLDTFQSDASPSHISMULFGR
BFLTKIYQLCPDPDLKALMLVRPGSMFIDNLSKESKFSDEGYSKRVYRVLVCEEDI
GLPKQFQHMWIIQNYPVNEVMEIKGGDHMAMLSDPKLCDCLSQISLKTA"
```

ORIGIN

|                        |        |               |      |  |  |
|------------------------|--------|---------------|------|--|--|
| Alignment Scores:      |        |               |      |  |  |
| Pred. No.:             | 66-64  | Length:       | 1077 |  |  |
| Score:                 | 793.00 | Matches:      | 152  |  |  |
| Percent Similarity:    | 73.36% | Conservative: | 38   |  |  |
| Best Local Similarity: | 58.69% | Mismatches:   | 67   |  |  |
| Query Match:           | 57.97% | Indels:       | 2    |  |  |
| DB:                    | 8      | Gaps:         | 1    |  |  |

US-10-780-002-2 (1-260) x AF269158 (1-1077)

Qy 3 GluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyLysGlyLysSerTrpTyLys 22  
Db 76 GAAGAAGAGCATTTTGTCTAGTTTCATGGAGTAAACCATGGAGCATGGTTGGTACAAA 135  
Qy 23 LeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAlaSer 42  
Db 136 CTGAAGAAGAGCTGGTGGCGGGGGTCCCGGGTGACGGCTGTGGACCTAGCCGCTCG 195  
Qy 43 GlyThrAspLeuArgLysIleGluLeuArgThrLeuTyAspTyThrLeuProLeu 62  
Db 196 GGCATCAACATCAGAGAAATTCAGGATGTGCACACATTCATGCATACAGTAGCCCTTG 255  
Qy 63 MetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeu 82  
Db 256 ATGGAGGTTTGGCATCATTCTCTCCGAGAAAGAGTCACTACTTGTGGACACAGTCTT 315  
Qy 83 GlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyProGlnLysIleTyAlaAlaVal 102  
Db 316 GGGGCGTCACCTTTGGCCCTTCCCGCGCACAAATTCACACAAATCTCCGTCGCTGTT 375  
Qy 103 PheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTy 122  
Db 376 TCTGTAACCTGCATTCATGCTGACACACACACCGCCCATCTTTTGTGTTTGGAGCATAT 435  
Qy 123 AsnGluArgThrProAlaGlu-----AsnTrpLeuAspThrGlnPheLeuProTyGly 140  
Db 436 TCTGAGAAGATGGGAAAGAGGACGACAGCTGGTTGGACACTCAATTTTCACAATGTGAC 495  
Qy 141 SerProGluGluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeu 160  
Db 496 CGGTCAATCCATCTCACATTTCCATGCTTTCCGCGCGAGTCTTGTGACTATCAAGATC 555  
Qy 161 TyrGlnLeuCySerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180  
Db 556 TATCAGCTTTGCTCTCTCCTGAGGATCTGGAGCTGGCAAGATGTTGGTGAGGCCAGATCA 615  
Qy 181 LeuPheMetGluAspLeuSerLysAlaLysTyPheThrAspGluArgPheGlySerVal 200  
Db 616 ATGTTTATAGCAACTTATCGAAGAAAGTAAAGTTTCAGCGATGAAGGATACCGATCTGTT 675  
Qy 201 LysArgValTyIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220  
Db 676 AAGCAGGTTTATCTGTATGCGAAGAGGATATTGCTCTCCCTAAGCAATTTTCAGCAGTGG 735  
Qy 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240  
Db 736 ATGATCCAAAACCTATCCCGTTAAATGAAGTGATGGAGATCAAGGGCGGTGATCAGATGCA 795  
Qy 241 MetLeuCyGluProGlnLysLeuCyAlaSerLeuLeuGluIleAlaHisLysTy 259  
Db 796 ATGCTTTCGAGCCCAAGAAAACCTTTGCGATTGTTCTGTCTCAGATTCTCTTAAAGTAT 852

```

RESULT 4
AX506625
LOCUS           792 bp      DNA      linear      PAT 27-SEP-2002
DEFINITION     Sequence 1320 from Patent WO216655.
ACCESSION     AX506625
VERSION       AX506625.1  GI:23387862
KEYWORDS
SOURCE        Arabidopsis thaliana (thale cress)
ORGANISM      Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS       Harper,J.F., Kreps,J., Wang,X. and Zhu,T.
TITLE         Stems-regulated genes of plants, transgenic plants containing
              same, and methods of use
JOURNAL       Patent: WO 0216655-A 1320 28-FEB-2002;
              The Scripps Research Institute (US); Syngenta Participations AG
              (CH)
FEATURES
             Location/Qualifiers
             1..792
               /organism="Arabidopsis thaliana"
               /mol_type="unassigned DNA"
               /db_xref="taxon:3702"
ORIGIN
Alignment Scores:
Pred. No.:      8,98e-62      Length:      792
Score:          768.00      Matches:     149
Percent Similarity: 71.04%      Conservative: 35
Best Local Similarity: 57.53%      Mismatches: 73
Query Match:    56.14%      Indels:      2
DB:             6           Gaps:          2

US-10-780-002-2 (1-260) x AX506625 (1-792)
QY  2  LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyClyTrpSerTrpTyr 21
Db  13  AAGAGGAAACAACATTTTGTACTAGTACATGTTGTCGTCACGGCGGTGGTGGTAC 72
QY  22  LysLeuLysProLeuLeuGluAlaGlyHisLysValThAlaLeuAspLeuAla 41
Db  73  AAGTTTAAGCCGCTGCTAGAGCGGTGGCCACCGCGTAACCTGCTGTGACCTTACTGCC 132
QY  42  SerGlyThrAsp---LeuArgLysLleGluLeuArgThLeuTyraSpTyrThrLeu 60
Db  133  TCGGAATAGACACACGAGGTGCTACTGACATCCCCACATCGAACCAATACTCGGAG 192
QY  61  ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValLleLeuValGlyHis 80
Db  193  CCAATTGACGAAGCTCTGACCTCATTTGCCAAATGATGAAGAAGTTGTGCTCGTGGTCC 252
QY  81  SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyProGlnLysLleTyraLa 100
Db  253  AGCTTTGGTGGCTTGAACCTTACCATAGCCATGAGCAAGAAAGTTTCCGAAAATCTCTGTC 312
QY  101  AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
Db  313  GCTGTATTCTTGACTGCTTTTCATGCGCGGACACCGAACACTCACCATCCTTCGCTTGGAC 372
QY  121  GlnTyraSerGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrgly 140
Db  373  AAGTTTGGGAAGCAACATGCTCAAGAAAGCATGGATGGGCACCGAATTCGAACCTTATGCT 432
QY  141  SerProGluGluProLeuThrSerMetPhePheGlyProLysPhePheLeuAlaHisLysLeu 160
Db  433  TCAGACAAATCCGGACTG---AGTATGTTTTTTAGCCCTGACTTCATGAAGTTGGGTCTC 489
QY  161  TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
Db  490  TACCAGCTTCTCCAGTTGAGGATCTTGAACCTGGGATTACTTTTAAATGAGGCCAGATCG 549
QY  181  LeuPheMetGluAspLeuSerLysAlaLysTyraPheThrAspGluArgPheGlySerVal 200

```

```

Db  550  TTATTTATTAACGATTATTCGAAGATGAAACATTCCTCGGATGAAGGATATGGGTCTGTT 609
QY  201  LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220
Db  610  CCTCGAGTTTTCATAGTGTGTAAGAGGACAAACAAATTCACAGAAACGCCAGAGATGG 669
QY  221  GlnIleAspAsnIleGlyValThrGluAlaIleGluLysGlyAlaAspHisMetAla 240
Db  670  ATGATTGATAATTTTCCGGTGAATTTAGTGTGAGATGGAGAGACAGATCATATGCCA 729
QY  241  MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyra 259
Db  730  ATGTTCTGCAAGCCTCAGCAACTCAGTGATTACTTCTGAAATATCGCGACAAATTC 786

RESULT 5
LOCUS           792 bp      mRNA      linear      PLN 05-JUN-2004
DEFINITION     Arabidopsis thaliana At2g23620 gene, complete cds.
ACCESSION     BT014881
VERSION       BT014881.1  GI:48310670
KEYWORDS      FLI CDNA.
SOURCE        Arabidopsis thaliana (thale cress)
ORGANISM      Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 792)
Cheuk,R., Chen,H., Kim,C.J., Shinn,P. and Ecker,J.R.
Arabidopsis ORF clones
Unpublished
2 (bases 1 to 792)
Cheuk,R., Chen,H., Kim,C.J., Shinn,P. and Ecker,J.R.
Submitted (05-JUN-2004) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
Location/Qualifiers
1..792
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="2"
/ecotype="Columbia"
/clone="U67545"
/notes="This clone is in pUNI 51"
1..792
/notes="unknown protein"
/codon_start=1
/product="At2g23620"
/protein_id="AA141864.1"
/db_xref="GI:48310671"
/translation="MSSEKRRKOHFVLVHSGCHGACWYKVKPLLEAVGHRVTAVDLAA
SGIDTTSTIDIPTEQYSEPLTKLLTSLPNDEKVLVGHSGFGLNLAIAMEKRPPEKI
SVAVFLTAPEMDTEHSFSLVLDKFGSNPQBAWMTPEPEPYGSDNSGLSMFSPDPMK
LGLQLSPDELDLGLLMPGSLFINDLSKKKNFSDGYSVPRVFIIVCKEDKAIFE
ERQRWMDINFPVNLVMEEMETDHPMFCKPOOLSDFLIKADKFY"
Alignment Scores:
Pred. No.:      8,98e-62      Length:      792
Score:          768.00      Matches:     149
Percent Similarity: 71.04%      Conservative: 35
Best Local Similarity: 57.53%      Mismatches: 73
Query Match:    56.14%      Indels:      2
DB:             8           Gaps:          2

US-10-780-002-2 (1-260) x BT014881 (1-792)
QY  2  LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyClyTrpSerTrpTyr 21
Db  13  AAGAGGAAACAACATTTTGTACTAGTACATGTTGTCGTCACGGCGGTGGTGGTAC 72

```

|   |  |
|---|--|
| Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A. |  |
| Direct Submission   |  |
| Submitted (15-JAN-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  |  |
| RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.   |  |
| The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.                   |  |
| Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.  |  |
| Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.   |  |
| Location/Qualifiers   |  |
| 1..935  |  |
| /organism="Arabidopsis thaliana"  |  |
| /mol_type="mRNA"  |  |
| /db_xref="taxon:3702"   |  |
| /chromosome="2"   |  |
| /clone="RAFL14-80-F10 (R20356)"   |  |
| /ecotype="Columbia"   |  |
| /note="This clone is in a modified pBluescript vector2 (lambda PS) as a BamHI/XhoI insert."   |  |
| gene  |  |
| 1..935  |  |
| /gene="At2g23620"   |  |
| 5'UTR   |  |
| 1..43   |  |
| /gene="At2g23620"   |  |
| CDS   |  |
| 44..835   |  |
| /gene="At2g23620"   |  |
| /codon_start=1  |  |
| /evidence="experimental"  |  |
| /product="putative acetone-cyanohydrin lyase"   |  |
| /protein_id="AAO22676.1"  |  |
| /db_xref="GI:27754457"  |  |
| /translation="MSEERKQHFVLVHSGCHGAWCYKVKPLLEAVGHRVAVDLAA SGVDTTSTIDPTCEQYSEPLTKLLTSLPNDKVLVGHSGFGLNLAIAMKFPFKKI SVAVFLTAFMDPTEHSPFVLDFKGSNNPQBAWMTPEPYGSDNSGLSMFSPFMK LGLYQLSPVDLELGLLMRPGSLFINDLSKMNKPSDRSGVSPRVFIVCKEDKAIPB BRQRWIDNFPVLVMEETDHPMFCKPQQLSDYFLKIADKFV"  |  |
| misc_difference   |  |
| 341   |  |
| /gene="At2g23620"   |  |
| /note="compared to genomic sequence resulting in an amino acid sequence difference"   |  |
| /replace="g"  |  |
| 836..935  |  |
| /gene="At2g23620"   |  |
| 3'UTR   |  |
| ORIGIN  |  |
| Alignment Scores:   |  |
| Pred. No.: 1.32e-61 Length: 935   |  |
| Score: 767.00 Matches: 149  |  |
| Percent Similarity: 71.04% Conservative: 35   |  |
| Best Local Similarity: 57.53% Mismatches: 73  |  |
| Query Match: 56.07% Indels: 2   |  |
| DB: 8 Gaps: 2   |  |



US-10-780-002-2 (1-260) x BT002859 (1-935)

```
QY 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTyr 21
Db 56 AAGAGGAACACATTTTGTACTAGTACATGGTTCGCCATGGCGCTGGTCTGGTAC 115
QY 22 LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla 41
Db 116 AAGGTTAAGCCGCTGTAGAGCGGTGGGCCACCGCGTAACCTGCTGGACTAGCTGCC 175
QY 42 SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu 60
Db 176 TCCGGAATAGACACACAGGCTCGATCACTGCACTCCCACTCCCACTGCGAACAATCTCGGAG 235
QY 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
Db 236 CCATTGACGAAGCTCTGACCTCAATTGCCAAATGATGAAGAGTTGTGCTGTTGTCAC 295
QY 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
Db 296 AGCTTTGGTGGCTTGAACCTAGCCATAGCCATGGAAAAGTTTCCCAAAAAATCTCTGTC 355
QY 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
Db 356 GTGTGATTCTTGACTGCTTTCATGCGGACACCGAACACTCACATCTCTGCTTGGAC 415
QY 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140
Db 416 AAGTTTGGAGCAACATGCTCAAGAGCATGATGGGCGCCAGCAATTCGAACCTTATGTT 475
QY 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
Db 476 TCAGACAATTCGGGACTG---AGTATGTTTTTTAGCCCTGACTTCATGAAGTTGGGTCTC 532
QY 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
Db 533 TACCAGCTTCTCCAGTTGAGGATCTTGAACTGGGATTACTTTTATGAGGCCAGATCG 592
QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
Db 593 TTATTATTATTAACGATTATTCGAAGATGAATAAACTCTCGGATGAAGGATATGGGTCTGT 652
QY 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220
Db 653 CCTCGAGTTTTCATAGTGTGTAAAGAGGCAAAAGCAATTCAGAGAAACGCCAGAGATGG 712
QY 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240
Db 713 ATGATTGATAAATTTCCGGTGAATTTAGTGTGGAGATGGAGGACAGATCATATGCCA 772
QY 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Db 773 ATGTTCTGCAAGCCTCAGCACTCAGTGATTACTTCTTGAAATTCGGGACAAATTC 829
```

```
RESULT 7
AY751530 891 bp mRNA linear PLN 11-OCT-2004
LOCUS
Catharanthus roseus protein S mRNA, complete cds.
DEFINITION
AY751530
ACCESSION
VERSION
AY751530.1 GI:53930669
```

KEYWORDS  
Catharanthus roseus (Madagascar periwinkle)

SOURCE  
Catharanthus roseus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons, core eudicots;  
asterids; lamids; Gentianales; Apocynaceae; Rauvolfioideae;  
Vinceae; Catharanthus.

REFERENCE  
1 (bases 1 to 891)

AUTHORS  
Lemenager, D., Rideau, M. and Clastre, M.  
TITLE  
Cloning of the Catharanthus roseus protein S (CrPS) associated with  
monoterpene indole alkaloid production  
JOURNAL  
unpublished  
REFERENCE  
2 (bases 1 to 891)

AUTHORS  
TITLE  
JOURNAL

Lemenager, D., Rideau, M. and Clastre, M.  
Direct Submission  
Submitted (15-SEP-2004) Laboratoire de Biologie Molculaire et  
Biochimie Vegetale, Faculte de Pharmacie, 31 Avenue Monge, Tours  
37200, France

FEATURES  
Location/Qualifiers

source 1..891  
/organism="Catharanthus roseus"  
/mol\_type="mRNA"  
/db\_xref="taxon:4058"  
48..824  
/note="hydrolase; CrPS; alpha/beta fold family protein"  
/codon\_start=1  
/product="protein S"  
/protein\_id="AAU5203.1"  
/db\_xref="GI:53830670"

CDS

48..824  
/note="hydrolase; CrPS; alpha/beta fold family protein"  
/codon\_start=1  
/product="protein S"  
/protein\_id="AAU5203.1"  
/db\_xref="GI:53830670"

/translation="MEYMKHFVTVHGVGHGAVVYKLPRIEAGHRCTAVNLAAGI  
NEKLEVRISIDYAAAPLLEVLDSVPENKIVILVHSGGNTAAVGMKFPNKISLAV  
FLNALPDTENRPSVLEEVYAKTPPEAWKOCFSAVGDPIITSLVCGPEFISSTLVH  
LSPEDHALGKILVRPSLFETDLKAEKTEEGFSGVPRVVVIAEDKTIPTPEFQW  
MIENNPVKEVKEIKGADHMPMPFKPDELSQLLDIAKHA"

ORIGIN

Alignment Scores:  
Pred. No.: 1.42e-58 Length: 891  
Score: 734.00 Matches: 144  
Percent Similarity: 70.66% Conservative: 39  
Best Local Similarity: 55.60% Mismatches: 74  
Query Match: 53.65% Indels: 2  
DB: 8 Gaps: 1

US-10-780-002-2 (1-260) x AY751530 (1-891)

```
QY 1 MetLysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTrp 20
Db 48 ATGGAAGTAAGTAAGCAATTTTGTGACGGTCCACGGCTAGGCCACGGAGCTTGGGTCTAT 107
QY 21 TyrLysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla 40
Db 108 TACAAGCTGAAACCTTAGAATTGAGGCAGACGCCACCGTTGTACGGCGTTTAATTTAGCT 167
QY 41 AlaSerGlyThrAspLeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu 60
Db 168 GCTTCTGGAATCAACGAAAGAACTCGAAGAGTTCGGTCTTCAATTGATTACGCCGCG 227
QY 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
Db 238 CGTTTGTGGAGGTTTGTGATTCTGTTCTCTGAAATGAGAAAGTAATACTCGTAGGCAT 287
QY 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
Db 288 AGCGCGGTGTATGACGCGCGCTGTGGGATGGAGAAATTTCCGAATAAGATCTCTTTA 347
QY 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
Db 348 GCTGTCTTTTGAATGCTATAATGCTGATACCGAAACACAGACCATCTTATGTTTGGAA 407
QY 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140
Db 408 GAGTACACTGCAAGACTCCACCAAGCATGGAAGGATTCGACATCTTCACGCTATGGA 467
QY 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
Db 468 GATCCCT-----CCTATTACATCTTTGTGCTGCGGTCTCTGAGTTTCATTTCACTCTG 521
QY 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
Db 522 TATCACCTCTCCCTTATTGAGGATCATGCATTTGGGAAAAAATATTTGGTAAGCCAGGGTCA 581
QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
Db 582 CTATTTATAGAGGATTTTATTGAAGCGCAGAAAAATTTACAGAGGAAGGATTTTGGATCAGTA 641
```





SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi  
1  
REFERENCE Harper,J.F., Kress,J., Wang,X. and Zhu,T.  
AUTHORS Stress-regulated genes of plants, transgenic plants containing  
TITLE same, and methods of use  
JOURNAL Patent: WO 0216655-A 1551 28-FEB-2002;  
The Scripps Research Institute (US); Syngenta Participations AG  
(CH)  
FEATURES Location/Qualifiers  
source  
1..792  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"  
ORIGIN  
Alignment Scores: Length: 792  
Pred. No.: 5,08e-55 Matches: 138  
Score: 695.00  
Percent Similarity: 67.18% Conservative: 36  
Best Local Similarity: 53.28% Mismatches: 83  
Query Match: 50.80% Indels: 2  
DB: 6 Gaps: 2  
US-10-780-002-2 (1-260) x AX506856 (1-792)  
Qy 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTrpTyr 21  
Db 13 AAGAGGAACACACTTCGTGCTAGTACATGGTGGCCGCGCATGGTCTGGTAC 72  
Qy 22 LysLeuLysProLeuLeuGluAlaGlyHisLysValThrAlaLeuAspLeuAla 41  
Db 73 AAGTTAAGCCCTTCTCGAGGCTTTGGGCCCATCGTGAACCGCTTAGACCTAGCTGCT 132  
Qy 42 SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu 60  
Db 133 TCCGGTATAGACACACACAGGTCATCTACATCTTCTACATGTGCAACAATATCTGAG 192  
Qy 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80  
Db 193 CCAATTGATCAGCTAATGACTTCAATTCGCCAATGATGAGAGGTTGTACTCGTTCGTCTAT 252  
Qy 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100  
Db 253 AGCTTTGGAGGTTTGAGTTTAGCTTAGCCATGGATAAGTTCCCGATAAAATCTCTGTC 312  
Qy 101 AlaValPheLeuAlaPheMetProAspSerValHisAsnSerPheValLeuGlu 120  
Db 313 TCTGTCTTCGTACTGCAATTCATCCCGACACCAACACTCACCATCGTTCGTCGAGGAA 372  
Qy 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140  
Db 373 AAGTTTGCAGCAGCATGACACAGAGGATGATGGCTCTGAGCTCGAGCATATGTT 432  
Qy 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160  
Db 433 TCAGATAATTCGGCTTG---TCTGTGTTCTTCAGCACCGCACTTCATGAAGCACCGTCTC 489  
Qy 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180  
Db 490 TACCAACTTCTCTCTGTGGAGGATCTTGAGCTTGGATTGCTTCTAAAGAGGCGCTAGTTCA 549  
Qy 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200  
Db 550 TTGTTTATTAAAGATTATCGAAGATGGAGAACTTTCTCGAAGAGGATATGACTGTT 609  
Qy 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220  
Db 610 CCTCGAGCTTACATTGTGTGCAAGAGGACACATTATCTCGGAGACCATCAACGATGG 669

Qy 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240  
Db 670 ATGATCATATATTTATTCGGCGAATTTAGTATTGAGATGGAGAGACTGATCATATGCCA 729  
Qy 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259  
Db 730 ATGTTTTCGAAACCTCAACTACTAAGTGACCATCTATTGGCAATCGTGCAATTTTC 786  
RESULT 10  
AY056068 792 bp mRNA linear PLN 30-SEP-2001  
LOCUS Arabidopsis thaliana At2g23600/f26B6.25 mRNA, complete cds.  
ACCESSION AY056068  
VERSION AY056068.1 GI:15810084  
KEYWORDS FLI\_CDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Arabidopsi  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi  
1 (bases 1 to 792)  
Koesema,E., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Shinn,P.,  
Banh,J., Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D.,  
Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A.,  
Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X.,  
Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J.,  
Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A.,  
Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,  
Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.  
Arabidopsis ORF clones  
Unpublished  
2 (bases 1 to 792)  
Koesema,E., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Shinn,P.,  
Banh,J., Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D.,  
Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A.,  
Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X.,  
Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J.,  
Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A.,  
Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,  
Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.  
Direct Submission  
Submitted (11-SEP-2001) Salk Institute Genomic Analysis Laboratory  
(SIGNAL), Plant Biology Laboratory, The Salk Institute for  
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,  
USA  
Riken Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN  
Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J.,  
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,  
Hayashizaki,Y. and Shinozaki,K.  
The Salk, Stanford, PGSC (SSP) Consortium members carried out the  
sequencing and annotation of the RAFL cDNAs: Koesema,E., Chen,H.,  
Cheuk,R., Kim,C.J., Meyers,M.C., Shinn,P., Banh,J., Bowser,L.,  
Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G.,  
Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,  
Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A.,  
Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,  
Davis,R.W., Theologis,A., and Ecker,J.R.  
Koesema,E. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally  
to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)  
contributed equally to this work as PIs.  
Location/Qualifiers  
1..792  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="U12118"  
/ecotype="Columbia"  
/note="This clone is in pUNI 51"  
1..792  
CDS

/notes="putative acetone-cyanohydrin lyase"

```
/codon_start=1
/product="At2g23600/F26B6.25"
/protein_id="AAL06968.1"
/db_xref="GI:15810085"
/translation="MSEERRKHQFVLVHGACHGAWCYKPKLLEALGHRVTALDLAA
SGIDTFRSTIDTSCQYSEPLQMLTSLPNDEKVLVGHSGGLSLALAMKFPDKI
SVSVFTAFMPDTHKSPFVEEKFASSMTPEKMGSELETYSGDSNGLSVFFSTDFMK
HRLYQLSPVEDLGLLKRPPSLFINELSKMENPSEKGVGSPRAYIVCKEDNIIE
DHQRWVHNYPANLVIMBETDHMPCKEQLLSHLLAIADNFC"
```

## ORIGIN

```
Alignment Scores:
Pred. No.: 5,08e-55 Length: 792
Score: 695.00 Matches: 138
Percent Similarity: 67.18% Conservative: 36
Best Local Similarity: 53.28% Mismatches: 83
Query Match: 50.80% Indels: 2
DB: 8 Gaps: 2

US-10-780-002-2 (1-260) x AY056068 (1-792)

Qy 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCyHisGlyGlyTrpSerTrpTyr 21
   |||
   ::|
Db 13 AAGAGGAAGCAACACTTCGTGCTAGTACATGGTGCCTGCCCGCATGGTCTGGTAC 72
   |||
Qy 22 LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAla 41
   |||
   ::|
Db 73 AAGTTAAGCTCTTCTCGAGGCTTTCGGCCATCGTGTACCCGCTAGACCTAGCTGCT 132
   |||
Qy 42 SerGlyThrAsp---LeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu 60
   |||
   ::|
Db 133 TCCGGTATAGACACCAACAGGTCAATCACTGACATTCCTACATGTGCAACATATCTGAG 192
   |||
Qy 61 ProLeuMetGluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
   |||
   ::|
Db 193 CCATTGATCGAGCTAAATGACTTCATTCGCCGAATGATGAGAAGGTGTACTCGTGTGTCAT 252
   |||
Qy 81 SerLeuGlyGlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
   |||
   ::|
Db 253 AGCTTGGAGGTTTGAGTTTACCTTACCATGGATGATGATTTCCCGATAAATCTCTGTC 312
   |||
Qy 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
   |||
   ::|
Db 313 TCTGTCTTCGTGATGCTCATTCATGCCCGACACCAACACTCACCATCGTTCGTGAGGAA 372
   |||
Qy 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140
   |||
   ::|
Db 373 AAGTTTGCAGCAGCATGACACCCAGGAAGGATGATGGCTCTGAGCTCGAGACATATGCT 432
   |||
Qy 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
   |||
   ::|
Db 433 TCAGATAATTCCGGCTTG---TCTGTGTTCTTCAGCACCAGCTTCATGAAGCACCCTCTC 489
   |||
Qy 161 TyrGlnLeuCySerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
   |||
   ::|
Db 490 TACCAACTTCTCTCTGGAGGATCTTGAGCTTGGATTGCTTCTAAGAGGCTAGTCTCA 549
   |||
Qy 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
   |||
   ::|
Db 550 TTGTTTTTAATGAATTCGAAAGATCGAGAACTTTTTCGAGAAAGGATGATGATCTGTT 609
   |||
Qy 201 LysArgValTyrIleValCyThrGluAspLysGlyIleProGluGluPheGlnArgTTP 220
   |||
   ::|
Db 610 CCTCGAGCTTACATTTGTGCAAGAGGACCAACATTTCTCGGAAGACCATCAACATGG 669
   |||
Qy 221 GlnIleAspAsnIleGlyValThrGluAlaIleLeuIleLysGlyAlaAspHisMetAla 240
   |||
   ::|
Db 670 ATGATCCATAATTTATCCGCGGAATTTAGTGAATGATGAGATGGAAGAGACTGATCATATGCCA 729
   |||
Qy 241 MetLeuCyGluProGlnLysLeuCyAlaSerLeuLeuGluIleAlaHisLysTyr 259
   |||
   ::|
Db 730 ATGTTTTGCAAACTCAACTACTAAGTGAACCATCTATTGGCAATCGCTGACAATTTC 786
```

RESULT 11  
AF361627  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AF361627 932 bp mRNA linear PLN 07-FEB-2002  
Arabidopsis thaliana At2g23600/F26B6.25 mRNA, complete cds.  
AF361627  
FLI CDNA.  
AF361627.1 GI:13605602  
Arabidopsis thaliana (thale cress)

Arabidopsis thaliana  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1 (bases 1 to 932)  
Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Banh,J., Bowser,L.,  
Carninci,P., Chung,M.K., Goldsmith,A.D., Hayashizaki,Y., Ishida,J.,  
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M.,  
Lin,J., Liu,S.X., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,  
Pham,P.K., Quach,H.L., Sakano,H., Sakurai,T., Satou,M., Seki,M.,  
Southwick,A., Toriumi,M., Yamada,K., Yu,G., Shinozaki,K.,  
Davis,R.W., Theologis,A. and Ecker,J.R.

Unpublished  
Arabidopsis cDNA clones  
2 (bases 1 to 932)  
Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Banh,J., Bowser,L.,  
Carninci,P., Chung,M.K., Goldsmith,A.D., Hayashizaki,Y., Ishida,J.,  
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M.,  
Lin,J., Liu,S.X., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,  
Pham,P.K., Quach,H.L., Sakano,H., Sakurai,T., Satou,M., Seki,M.,  
Southwick,A., Toriumi,M., Yamada,K., Yu,G., Shinozaki,K.,  
Davis,R.W., Theologis,A. and Ecker,J.R.

Direct Submission  
Submitted (15-MAR-2001) Salk Institute Genomic Analysis Laboratory  
(SIGnAL), Plant Biology Laboratory, The Salk Institute for  
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,  
USA  
RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN  
Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J.,  
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,  
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the  
sequencing and annotation of the RAFL cDNAs: Shinn,P., Chen,H.,  
Cheuk,R., Kim,C.J., Koseema,E., Meyers,M.C., Tracy,S.E., Banh,J.,  
Bowser,L., Chung,M.K., Goldsmith,A.D., Jones,T., Karlin-Neumann,G.,  
Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,  
Palm,C.J., Pham,P.K., Quach,H.L., Sakano,H., Southwick,A.,  
Tang,C.C., Toriumi,M., Yamada,K., Yu,G., Davis,R.W., Theologis,A.,  
and Ecker,J.R.

Shinn,P. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to  
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)  
contributed equally to this work as PIs.

Location/Qualifiers  
1..932  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="RAFL09-12-L06 (R12118)"  
/ecotype="Columbia"  
1..13  
14..805  
/note="putative acetone-cyanohydrin lyase"  
/codon\_start=1  
/product="At2g23600/F26B6.25"  
/protein\_id="AAK32795.1"  
/db\_xref="GI:13605603"  
/translation="MSEERRKHQFVLVHGACHGAWCYKPKLLEALGHRVTALDLAA  
SGIDTFRSTIDTSCQYSEPLQMLTSLPNDEKVLVGHSGGLSLALAMKFPDKI  
SVSVFTAFMPDTHKSPFVEEKFASSMTPEKMGSELETYSGDSNGLSVFFSTDFMK  
HRLYQLSPVEDLGLLKRPPSLFINELSKMENPSEKGVGSPRAYIVCKEDNIIE

Shinn,P. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to  
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)  
contributed equally to this work as PIs.

FEATURES  
source

5'UTR  
CDS

```
3'UTR
ORIGIN
DHORWMIHNPANLVIEEMETDMPFCKPQLLSHLLAIADNFC"
806..932

Alignment Scores:
Pred. No.: 6,01e-55 Length: 932
Score: 695.00 Matches: 138
Percent Similarity: 67.18% Conservative: 36
Best Local Similarity: 53.28% Mismatches: 83
Query Match: 50.80% Indels: 2
DB: Gaps: 2

US-10-780-002-2 (1-260) x AF361627 (1-932)

QY 2 LysGluGlyLysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyr 21
Db 26 NAGAGGAGCAACACTTCGTGCTAGTACATGCTGCGCCAGCGCATGGTGGTGTAC 85

QY 22 LysLeuLysProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAla 41
Db 86 AAGGTTAAGCCTCTTCGAGGCTTTGGGCCATCGTGTAAACCGCCTTAGACCTAGCTGCT 145

QY 42 SerGlyThrAsp--LeuArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeu 60
Db 146 TCGGTATAGACACACACAGGTCATCACTGACATCTTACATGTGAACATATTTCTGAG 205

QY 61 ProLeuMetGluMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHis 80
Db 206 CCATTGATGACGTAATGACTTCATTGCCGAATGATGAGAGGTTGTACTCGTTGGTCA 265

QY 81 SerLeuGlyGlyMetAsnLeuLeuAlaMetGluLysTyrProGlnLysIleTyrAla 100
Db 266 AGCTTTGGAGGTTTGAGTTAGCTTAGCCATGGATAAGTTTCCCGATAAAATCTCTGC 325

QY 101 AlaValPheLeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGlu 120
Db 326 TCTGCTTCGTGACTGATTCATGCCGACATTCGCCACCAACACTCACCTGTTCTGCGGAA 385

QY 121 GlnTyrAsnGluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGly 140
Db 386 AAGTTTGCAGCAGCATGACACAGAGGATGGATGGGCTCTGAGCTCGAGACATATGCT 445

QY 141 SerProGluGluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeu 160
Db 446 TCAGATAATTCGGCTTG--TCTGTGTTCTTCAGCACCGACTTCATGAAGCACCGCTCTC 502

QY 161 TyrGlnLeuCysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSer 180
Db 503 TACCAACTTTCCTCTGGAGGATCTTGAGCTTGGATTGCTTCTAAAGAGGCTTAGTTCA 562

QY 181 LeuPheMetGluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerVal 200
Db 563 TTGTTTATTAAATTAATTCGAAGATGGAGAACTTCTCGAGAAAGGGTATGGATCTGTT 622

QY 201 LysArgValTyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrp 220
Db 623 CCTCGAGCTTACATTGTGCAAGAGGACCAATTCATCTCGAAGACCATCAAGCATGG 682

QY 221 GlnIleAspAsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAla 240
Db 683 ATGATCCATTAATATCCGCGGAATTTAGTGATTGATGGAGAGAGACTGATCATATGCCA 742

QY 241 MetLeuCysGluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Db 743 ATGTTTGCACAACTCACTACTACTAGTGACCATCTATTGGCAATCGCTGACAATTC 799

RESULT 12
CQ806104
LOCUS
DEFINITION
Sequence 2515 from Patent WO2004035798.
CQ806104
ACCESSION
CQ806104.1 GI:47111686
KEYWORDS
```

```
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
1 Inze, D., de Veylder, L. and Vlieghe, K.
AUTHORS
Identification of novel e2f target genes and use thereof
JOURNAL
Patent: WO 2004035798-A 2515 29-APR-2004;
CropDesign N.V. (BE)
FEATURES
Location/Qualifiers
source
1..792
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Alignment Scores:
Pred. No.: 1.47e-54 Length: 792
Score: 690.00 Matches: 134
Percent Similarity: 69.92% Conservative: 45
Best Local Similarity: 52.34% Mismatches: 75
Query Match: 50.44% Indels: 2
DB: Gaps: 2

US-10-780-002-2 (1-260) x CQ806104 (1-792)

QY 5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTrpSerTrpTyrLysLeuLys 24
Db 22 CAACACGTCGTTCTAGTACATGCTTGCCTCCATCGCGCTGCTGCTGCAAGGTTAAG 81

QY 25 ProLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAspSerGlyThr 44
Db 82 CCGAGCTCGAGGCTTCTGCCACCGGTAACCGGCTAGATCTAGCTGCCCTCGGTATA 141

QY 45 AspLeu--ArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMet 63
Db 142 GACATGACCAAGTCATCATCAGATATATCCATCGCAACATATCTCAGAGCCATTGATG 201

QY 64 GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 83
Db 202 CAGCTAATGACTCACTACCATGATGATGAGAGGTTGTGCTTGTGCTAGCTTAGGA 261

QY 84 GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPhe 103
Db 262 GGTTCGATTGCTATGCTATGGCCATGATGTTTCCGACCAAAATCTCTGTTCTGCTTT 321

QY 104 LeuAlaAlaPheMetProAspSerValHisAsnSerPheValLeuGluGlnTyrAsn 123
Db 322 GTGACTGTATGATGCCAGACACCAACACTCACCATCTCTCGTATGGGATAAGCTAAGA 381

QY 124 GluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGlu 143
Db 382 AAGAAACTTTCACGAGAGGAATGGTTAGACACCGTGTTCAGGAC--GAGAAACCTGAT 438

QY 144 GluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeu 163
Db 439 TTTCTACGAGTTCGATTTTGGACCAAGATTCATGGCCCAAGAACTTGTATCAGTTG 498

QY 164 CysSerProGluAspLeuAlaSerSerLeuValArgProSerSerLeuPheMet 183
Db 499 TCTCCAGTCCAAGATCTTGAATTCGGCAAAATTTGTTGAGGGCAACCCATTGATTAAAG 558

QY 184 GluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgVal 203
Db 559 AAGATATGCGCAGAGAGAGAGGTTCTAGTGAGGAAGGATACGGATCCGCTTACAGCTATA 618

QY 204 TyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAsp 223
Db 619 TTTATTGTATGCGGAAGAGGATCTTGTGTCACCCGGAAGATTCACCGCATCGATGATCAGC 678

QY 224 AsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243
```

```

Db      679  AACTTTCCCAAAAGAAAGTAATGGAGATCAAAAGCAGCATATGCGCAATGTTCTCC 738
QY      244  GluProGlnLysLeuCysAlaSerLeuLeuGluLeuAlaHisLysTyr 259
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      739  AAGCTCAACAACACTATGCTCTCTCTTGAGATTGCAATATAATAT 786

RESULT 13
LOCUS   AX505522                               792 bp    DNA    linear    PAT 27-SEP-2002
DEFINITION Sequence 217 from Patent WO0216655.
ACCESSION AX505522
VERSION   AX505522.1  GI:23386759
KEYWORDS
SOURCE   Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS  Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
TITLE    Stress-regulated genes of plants, transgenic plants containing
          same, and methods of use
JOURNAL  Patent: WO 0216655-A 217 28-FEB-2002;
          The Scripps Research Institute (US) ; Syngenta Participations AG
          (CH)
FEATURES             Location/Qualifiers
     source          1..792
                     /organism="Arabidopsis thaliana"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3702"
ORIGIN
Alignment Scores:
Pred. No.:          1,478-54      Length:          792
Score:              690.00        Matches:         134
Percent Similarity: 69.92%        Conservative:    45
Best Local Similarity: 52.34%      Mismatches:     75
Query Match:        50.44%        Indels:         2
DB:                  6            Gaps:             2

US-10-780-002-2 (1-260) x AX505522 (1-792)

QY      5   LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrSerTyrLysLeuLys 24
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      22  CAACAGCTGCTTCTAGTACATGGTGTGCCATGGCGCTGCTGGTACAGGTAG 81

QY      25  ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaSerGlyThr 44
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      82  CCGCAGCTCGAGGCTTCTGGCCACCGGTAAACCGCGTAGATCTAGCTGCTCCGGTATA 141

QY      45  AspLeu---ArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMet 63
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      142 GACATGACCCAGGTCATACAGATATATCCATATATCCATGCGAACAATACTCAGAGCCATTGATG 201

QY      64  GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 83
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      202 CAGCTAATGACCTCACTACAGATGATGAGAGAGTTGCTGTTGTTGTCATGCTAGGA 261

QY      84  GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPhe 103
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      262 GGTTCAGTTTACTATGCGCATGGATATGTTTCCGACCAAAATCTCTGTTCTGCTTT 321

QY      104 LeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsn 123
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      322 GTGACTGCTATGATGTCAGACACCAACAACACTCACCATCTCTGATGGGATAGCTAAGA 381

QY      124 GluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGlu 143
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      382 AAGAAACTTCACGAGAGGATGGTTAGACCGTGTTTACGAGC---GAGAAACCTGAT 438

QY      144 GluProLeuThrSerMetPhePheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeu 163
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      439 TTTCCTAGCGAGTTTGGATTTTTGGACCAAGATTATCGCCCAAGAACTTGTATCAGTTG 498

```

```

QY      164  CysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMet 183
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      499  TCTCCAGTCAGAGATCTTGAATTTGGCAAAATGTTGGTGGGCAACCCATTTGATTAAG 558

QY      184  GluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgVal 203
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      559  AAGATATATGCGACAGAGAAGAGCTTCAGTGAGGAAGGATACGGATCCGTTACACGTATA 618

QY      204  TyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAsp 223
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      619  TTTATTGTATGCGAAAGAGATCTTGTGTCCACCGAAGATTACACGAGTCGATGATCAGC 678

QY      224  AsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      679  AACTTTCCCAAAAGAGTAAATGAGATCAAAAGCGCAGATCATATGCCAATGTTCTCC 738

QY      244  GluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      739  AAGCCTCAACAACACTATGCTCTCTCTTGGAGATTGCAATAAATAT 786

RESULT 14
LOCUS   BT015031                               792 bp    mRNA    linear    PLN 10-JUL-2004
DEFINITION Arabidopsis thaliana At2g23610 mRNA, complete cds.
ACCESSION BT015031
VERSION   BT015031.1  GI:50198957
KEYWORDS  FLI-CDNA.
SOURCE    Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 792)
AUTHORS  Shinn, P., Chen, H., Cheuk, R., Kim, C.J. and Ecker, J.R.
TITLE    Arabidopsis ORF clones
JOURNAL  Unpublished
REFERENCE 2 (bases 1 to 792)
AUTHORS  Shinn, P., Chen, H., Cheuk, R., Kim, C.J. and Ecker, J.R.
TITLE    Direct Submission
JOURNAL  Submitted (10-JUL-2004) Salk Institute Genomic Analysis Laboratory
          (SIGNAL), Plant Biology Laboratory, The Salk Institute for
          Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
          USA
FEATURES             Location/Qualifiers
     source          1..792
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="2"
                     /clone="U67063"
                     /ecotype="Columbia"
                     /note="This clone is in pUNI 51"
     CDS              1..792
                     /note="unknown protein"
                     /codon_start=1
                     /product="At2g23610"
                     /protein_id="AA70482.1"
                     /db_xref="GI:50198958"
                     /translation="MSEERKQHVVLVHGACHGAWCYKVKPOLSEASGHRVTVADLAA
                     SGIDWTRSTIDISTCEQYSEPLMOLMTSLPDKEKVLVGHSGISGLAMMDMPPTKI
                     SVSVPTAMPDPTKHSFVWDKLRKETSREWLDTVETSEKPDPSFEWIFGPBFMA
                     KNLVQLSPQDLLEAKMLVRANPLIKDMAERSFSEGYSGVTRIFIVCGKDLVSPE
                     DYQRSMISNFPFPEYMEIKDADHMPFMSKPOLCALLLEIAANKYA"
ORIGIN
Alignment Scores:
Pred. No.:          1,478-54      Length:          792
Score:              690.00        Matches:         134
Percent Similarity: 69.92%        Conservative:    45
Best Local Similarity: 52.34%      Mismatches:     75
Query Match:        50.44%        Indels:         2
DB:                  8            Gaps:             2

```

US-10-780-002-2 (1-260) x BT015031 (1-792)

```
Qy 5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrPsrTyrTyrLysLeuLys 24
Db 22 CAACACGTCGTTCTAGTACATGGTCTTGGCCATGGCGCTGGTGTGGTACAGGTTAAG 81
Qy 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAlaSerGlyThr 44
Db 82 CGCAGCTCGAGGCTTCTGGCCACGCGTAAACGCGTAGATCTAGCTGCCTCCGCTATA 141
Qy 45 AspLeu--ArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMet 63
Db 142 GACATGACGACGATCAATCAGATATATCCACGCAACAATCTCAGACGATTGATG 201
Qy 64 GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 83
Db 202 CAGCTAATGACCTCACTACCAGATGATGAGAAGGTTGTCTTGTGGTACATGCTTAGGA 261
Qy 84 GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPhe 103
Db 262 GGTGTTGAGTTAGTATGATGATGATGATGTTTCCGACCAAAATCTCTGTCTCTCTT 321
Qy 104 LeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsn 123
Db 322 GTGACTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 381
Qy 124 GluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGlu 143
Db 382 AAAGAACTTCCAGAGAGAAATGGTTAGACACGGTGTTCAGAGC--GAGAACTGAT 438
Qy 144 GluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeu 163
Db 439 TTTCTAGGAGGTTTGGATTTTGGACCAAGAATTCATGCCAAGAATCTTGTATCAGTTG 498
Qy 164 CysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMet 183
Db 499 TCTCCAGTCACAGATCTTGAATGGCGAAAATGTTGGTGGAGGCAAAACCATGATTAAG 558
Qy 184 GluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgVal 203
Db 559 AAAGATATGCGAGAGAGAGAGCTTCAGTGAGGAGGATACGGATCCGTTACAGGTATA 618
Qy 204 TyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAsp 223
Db 619 TTTATTTGATGCGGAAAGATCTTGTGTCAACCGAAGATTACAGCGATCGATGATCAGC 678
Qy 224 AsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243
Db 679 AACTTTCCCAAAAGAGTAATGGAGATCAAGAGCGCAGATCATATGCCAATGTTCTCC 738
Qy 244 GluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Db 739 AAGCCTCAACAATATGCTCTTCTCTTGGAGATTGCAATAATAT 786
```

## RESULT 15

BT014901 878 bp mRNA linear PLN 11-JUN-2004  
LOCUS Arabidopsis thaliana At2g23610 gene, complete cds.

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

TITLE Direct Submission  
JOURNAL Submitted (11-JUN-2004) Salk Institute Genomic Analysis Laboratory (Signal), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

FEATURES  
source

Location/Qualifiers  
1..878  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="S67063"  
/ecotype="Columbia"  
/note="This clone is in pUNI 51"  
8..799  
/note="unknown protein"  
/codon\_start=1  
/product="At2g23610"  
/protein\_id="AA046030.1"  
/db\_xref="GI:48596979"  
/translation="MSERRKQHVVLVHGACHGAWCYKVKPQLEASGHRVTAVDLAA  
SGIDMTSTIDTISTCEQYSBPLMQLMTSLPDEKVLVHSLGSLSLAMAMDPFTKI  
SVSVFTAMSPDTRKHSFVMDKLKETSREWLDTFTSEKPDFFSEFWLFGPEFMA  
KNLYQLSPVODLELAKMLVRANPLIKDMAERRSFSEGVSTRIFTVCQGLVSP  
DYQSRMISNPPEKVMKIDADHMPFSKPQQLCALLETANKYA"

## CDS

## ORIGIN

Alignment Scores:  
Pred. No.: 1.64e-54 Length: 878  
Score: 690.00 Matches: 134  
Percent Similarity: 69.92% Conservative: 45  
Best Local Similarity: 52.34% Mismatches: 75  
Query Match: 50.44% Indels: 2  
DB: 8 Gaps: 2

US-10-780-002-2 (1-260) x BT014901 (1-878)

```
Qy 5 LysHisPheValLeuValHisGlyAlaCysHisGlyGlyTyrPsrTyrTyrLysLeuLys 24
Db 29 CAACACGTCGTTCTAGTACATGGTCTTGGCCATGGCGCTGGTGTGGTACAGGTTAAG 88
Qy 25 ProLeuLeuGluAlaAlaGlyHisLysValThrAlaLeuAspLeuAlaAlaSerGlyThr 44
Db 89 CGCAGCTCGAGGCTTCTGGCCACCGGTAAACCGCGTAGATCTAGCTGCCTCCGCTATA 148
Qy 45 AspLeu--ArgLysIleGluGluLeuArgThrLeuTyrAspTyrThrLeuProLeuMet 63
Db 149 GACATGACGAGTCAATCACAGATATATCCATGCGAACAATATCTCAGAGCCATTGATG 208
Qy 64 GluLeuMetGluSerLeuSerAlaAspGluLysValIleLeuValGlyHisSerLeuGly 83
Db 209 CAGCTAATGACCTCACTACAGATGATGAGAAGGTTGTCTTGTGGTACATGCTAGGA 268
Qy 84 GlyMetAsnLeuGlyLeuAlaMetGluLysTyrProGlnLysIleTyrAlaAlaValPhe 103
Db 269 GGTGAGTTAGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 328
Qy 104 LeuAlaAlaPheMetProAspSerValHisAsnSerSerPheValLeuGluGlnTyrAsn 123
Db 329 GTGACTGCTATGATGCCAGACACCAACACCTCACCATCTCTCGTATGGGATAAGTAAGA 388
Qy 124 GluArgThrProAlaGluAsnTrpLeuAspThrGlnPheLeuProTyrGlySerProGlu 143
Db 389 AAAGAACTTCCAGAGAGAAATGGTTAGACACCGGTTTACGAGC--GAGAAACCTGAT 445
Qy 144 GluProLeuThrSerMetPheGlyProLysPheLeuAlaHisLysLeuTyrGlnLeu 163
Db 446 TTTCTAGCGAGTTTGGATTTTGGACCAAGAATTCATGCCAAGAATCTTGTATCAGTTG 505
Qy 164 CysSerProGluAspLeuAlaLeuAlaSerSerLeuValArgProSerSerLeuPheMet 183
Db 506 TCTCCAGTCCAAGATCTTGAATTGGCAAAATGTTGGTGGAGGCAAAACCATGATTAAG 565
```

```
Qy 184 GluAspLeuSerLysAlaLysTyrPheThrAspGluArgPheGlySerValLysArgVal 203
Db ::::::::::: :::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
566 AAGATATGGCAGAGAGAGAGCTTCAGTGAGGAGGATACGGATCCGTTACACGTATA 625
Qy 204 TyrIleValCysThrGluAspLysGlyIleProGluGluPheGlnArgTrpGlnIleAsp 223
Db ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
626 TTTATTGTATCGCGAAGAGGATCTTGTGTACCCGAGATTACCAGCGATCGATCATCAGC 685
Qy 224 AsnIleGlyValThrGluAlaIleGluIleLysGlyAlaAspHisMetAlaMetLeuCys 243
Db ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
686 AACTTCCCCCAAGAGAGTAAATGGAGATCAAGACGAGATCATATGCCAATGTTCTTCC 745
Qy 244 GluProGlnLysLeuCysAlaSerLeuLeuGluIleAlaHisLysTyr 259
Db ::::::::::: ::::::::::: ::::::::::: ::::::::::: :::::::::::
746 AAGCCTCAACAACATATGTGCTTCTCTTGGAGATTGCAATAAATAT 793
```

Search completed: September 12, 2005, 21:55:34  
Job time : 3435 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 15:37:37 ; Search time 733.243 Seconds  
(without alignments)  
9663.988 Million cell updates/sec

Title: US-10-780-002-1  
Perfect score: 1079  
Sequence: 1 acgcgggaagaagaagaaaa.....gaaaaaaaaaaaaaaaaa 1079

Scoring table: OLIGO\_NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 7351250 seqs, 3283620254 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 14702500

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:\*  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:\*  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:\*  
22: /cgn2\_6/ptodata/1/pubpna/US10I\_NEW\_PUB.seq:\*  
23: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 1079  | 100.0       | 1079   | 21 | US-10-780-002-1   |
| 2          | 29    | 2.7         | 607    | 13 | Sequence 1, Appli |
| 3          | 29    | 2.7         | 607    | 17 | Sequence 241778,  |
| 4          | 28    | 2.6         | 615    | 20 | Sequence 241778,  |
| 5          | 27    | 2.5         | 207    | 17 | Sequence 17538, A |
| 6          | 27    | 2.5         | 207    | 18 | Sequence 17538, A |
| 7          | 27    | 2.5         | 321    | 9  | Sequence 4280, Ap |

|   |    |    |     |       |    |                      |                   |
|---|----|----|-----|-------|----|----------------------|-------------------|
| c | 8  | 27 | 2.5 | 446   | 10 | US-09-873-367C-946   | Sequence 946, App |
| c | 9  | 27 | 2.5 | 446   | 21 | US-10-843-641A-946   | Sequence 946, App |
| c | 10 | 27 | 2.5 | 511   | 20 | US-10-425-115-102591 | Sequence 102591,  |
| c | 11 | 27 | 2.5 | 1795  | 9  | US-09-925-300-754    | Sequence 754, App |
| c | 12 | 27 | 2.5 | 1795  | 17 | US-10-264-237-1253   | Sequence 1253, Ap |
| c | 13 | 26 | 2.4 | 332   | 17 | US-10-242-535A-54828 | Sequence 54828, A |
| c | 14 | 26 | 2.4 | 332   | 18 | US-10-085-783A-54828 | Sequence 54828, A |
| c | 15 | 26 | 2.4 | 379   | 19 | US-10-437-963-72087  | Sequence 72087, A |
| c | 16 | 26 | 2.4 | 406   | 20 | US-10-425-115-111034 | Sequence 111034,  |
| c | 17 | 26 | 2.4 | 814   | 9  | US-09-780-717-34     | Sequence 34, Appl |
| c | 18 | 26 | 2.4 | 1099  | 22 | US-10-487-092-68     | Sequence 68, Appl |
| c | 19 | 26 | 2.4 | 98345 | 21 | US-10-461-862-136    | Sequence 136, App |
| c | 20 | 25 | 2.3 | 135   | 17 | US-10-242-535A-3360  | Sequence 3360, Ap |
| c | 21 | 25 | 2.3 | 135   | 18 | US-10-085-783A-3360  | Sequence 3360, Ap |
| c | 22 | 25 | 2.3 | 148   | 20 | US-10-357-930-60780  | Sequence 60780, A |
| c | 23 | 25 | 2.3 | 157   | 17 | US-10-242-535A-17475 | Sequence 17475, A |
| c | 24 | 25 | 2.3 | 157   | 18 | US-10-085-783A-17475 | Sequence 17475, A |
| c | 25 | 25 | 2.3 | 158   | 20 | US-10-357-930-61029  | Sequence 61029, A |
| c | 26 | 25 | 2.3 | 172   | 20 | US-10-425-115-166206 | Sequence 166206,  |
| c | 27 | 25 | 2.3 | 182   | 20 | US-10-425-115-85062  | Sequence 85062, A |
| c | 28 | 25 | 2.3 | 194   | 20 | US-10-357-930-60898  | Sequence 60898, A |
| c | 29 | 25 | 2.3 | 214   | 9  | US-09-983-965-591    | Sequence 591, App |
| c | 30 | 25 | 2.3 | 219   | 9  | US-09-764-846-51     | Sequence 51, Appl |
| c | 31 | 25 | 2.3 | 219   | 14 | US-10-081-483-51     | Sequence 51, Appl |
| c | 32 | 25 | 2.3 | 221   | 20 | US-10-357-930-61823  | Sequence 61823, A |
| c | 33 | 25 | 2.3 | 257   | 20 | US-10-425-115-155173 | Sequence 155173,  |
| c | 34 | 25 | 2.3 | 258   | 18 | US-10-424-599-123984 | Sequence 123984,  |
| c | 35 | 25 | 2.3 | 276   | 20 | US-10-425-115-114837 | Sequence 114837,  |
| c | 36 | 25 | 2.3 | 282   | 14 | US-10-066-543-2060   | Sequence 2060, Ap |
| c | 37 | 25 | 2.3 | 290   | 20 | US-10-357-930-61438  | Sequence 61438, A |
| c | 38 | 25 | 2.3 | 298   | 19 | US-10-437-963-90281  | Sequence 90281, A |
| c | 39 | 25 | 2.3 | 310   | 20 | US-10-723-860-1322   | Sequence 1322, Ap |
| c | 40 | 25 | 2.3 | 319   | 18 | US-10-424-599-73498  | Sequence 73498, A |
| c | 41 | 25 | 2.3 | 324   | 20 | US-10-425-115-135371 | Sequence 135371,  |
| c | 42 | 25 | 2.3 | 341   | 10 | US-09-814-353-4668   | Sequence 4668, Ap |
| c | 43 | 25 | 2.3 | 341   | 20 | US-09-814-353-10967  | Sequence 10967, A |
| c | 44 | 25 | 2.3 | 341   | 20 | US-10-425-115-142143 | Sequence 142143,  |
| c | 45 | 25 | 2.3 | 356   | 18 | US-10-424-599-77760  | Sequence 77760, A |

ALIGNMENTS

RESULT 1  
US-10-780-002-1  
; Sequence 1, Application US/107800002  
; Publication No. US20050034196A1  
; GENERAL INFORMATION:  
; APPLICANT: Kumar, Dhirenda  
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN  
; FILE OF INVENTION: ENCODING NUCLEIC ACIDS, SAB2, AND METHODS OF USE THEREOF  
; FILE REFERENCE: 3670-PO2652WO  
; CURRENT APPLICATION NUMBER: US/10/780,002  
; CURRENT FILING DATE: 2004-02-17  
; PRIOR APPLICATION NUMBER: PCT/US02/26312  
; PRIOR FILING DATE: 2002-08-16  
; PRIOR APPLICATION NUMBER: 60/312,863  
; PRIOR FILING DATE: 2001-08-16  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Nicotiana tabacum  
US-10-780-002-1

Query Match 100.0%; Score 1079; DB 21; Length 1079;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1079; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACGCCGGGAAAGAAAGAAACTTAAACAGGCATATAATTCAAATGAAGAGAAACACT 60  
|||||

Db 1 ACGCGGGGAAAGAAAGAACTAAACAGGCATATAAATTCAAATGAAGGAGGAAACACT 60  
Qy 61 TTGTTTTAGTACATGGTGCATCCCATGGAGGTTGGAGTTGGTACAAAGCTAAAGCCACTGC 120  
Db 61 TTGTTTTAGTACATGGTGCATCCCATGGAGGTTGGAGTTGGTACAAAGCTAAAGCCACTGC 120  
Qy 121 TAGAAGCTGCAGGCCATAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCAGCTGATTTGA 180  
Db 121 TAGAAGCTGCAGGCCATAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCAGCTGATTTGA 180  
Qy 181 GAAATAATAGAGAGCTTGCACACTTTATGATTATACCTTTGGCATTTGATGGATTTGATGG 240  
Db 181 GAAATAATAGAGAGCTTGCACACTTTATGATTATACCTTTGGCATTTGATGGATTTGATGG 240  
Qy 241 AATCTCTTTACGACAGATGAGAGGTTATATTAGTGGGCATAGTCTTGGTGGTATGAATT 300  
Db 241 AATCTCTTTACGACAGATGAGAGGTTATATTAGTGGGCATAGTCTTGGTGGTATGAATT 300  
Qy 301 TGGGACTTGTCTATGGAAGATATCCAAAGATCTATGCTGCTGTTCTTCTGGCTGCTT 360  
Db 301 TGGGACTTGTCTATGGAAGATATCCAAAGATCTATGCTGCTGTTCTTCTGGCTGCTT 360  
Qy 361 TCATGCTGATTTCTGTTCAACACTCTCTCTTTGTTTGGAAACAGTATATAGCGGAGCG 420  
Db 361 TCATGCTGATTTCTGTTCAACACTCTCTCTTTGTTTGGAAACAGTATATAGCGGAGCG 420  
Qy 421 CAGCGGAGAAATGGTTGGATATCTAGTTTTTACCATATGTTTCCCTGAAAGCCACTGA 480  
Db 421 CAGCGGAGAAATGGTTGGATATCTAGTTTTTACCATATGTTTCCCTGAAAGCCACTGA 480  
Qy 481 CATCCATGTTTTTGGCCCAAAAGTTCTTGGCTCACAGCTCTACAGCTATGCTCTCTG 540  
Db 481 CATCCATGTTTTTGGCCCAAAAGTTCTTGGCTCACAGCTCTACAGCTATGCTCTCTG 540  
Qy 541 AGGATCTTCATTTAGCATCATCATGTTGAGACCAAGCTCTTTGTTTATGGAAGCACTAT 600  
Db 541 AGGATCTTCATTTAGCATCATCATGTTGAGACCAAGCTCTTTGTTTATGGAAGCACTAT 600  
Qy 601 CGAAGGCCAAGTATTTTCAGATGAACGGTTTGGATCAGTGAAGAGAGTTTACATGTGT 660  
Db 601 CGAAGGCCAAGTATTTTCAGATGAACGGTTTGGATCAGTGAAGAGAGTTTACATGTGT 660  
Qy 661 GCATCTGAGTAAAGCATACAGAGATTCACGGATGCGCAATTCGCAATTCACACATTTGGT 720  
Db 661 GCATCTGAGTAAAGCATACAGAGATTCACGGATGCGCAATTCGCAATTCACACATTTGGT 720  
Qy 721 TCACCTGAAGCAATAGAGATTAAGGTCGTGATCAGATGCAATGCTATCGAGCCCAAA 780  
Db 721 TCACCTGAAGCAATAGAGATTAAGGTCGTGATCAGATGCAATGCTATCGAGCCCAAA 780  
Qy 781 AACTTTGGCCCTCTCTCTTGGAAATGGCCATAAATACAACTGATCTTACATTTATGCT 840  
Db 781 AACTTTGGCCCTCTCTCTTGGAAATGGCCATAAATACAACTGATCTTACATTTATGCT 840  
Qy 841 TCGTCTCATGTCAAGATTTTCAGTGCATGCTGTAATTTTTTCTATTTTTCGACCGGCG 900  
Db 841 TCGTCTCATGTCAAGATTTTCAGTGCATGCTGTAATTTTTTCTATTTTTCGACCGGCG 900  
Qy 901 ATAACCTGCTTTGCTCTATTTAAGGATTCAGTAATTTTCACCTCTCTAGTGTGGAAGCT 960  
Db 901 ATAACCTGCTTTGCTCTATTTAAGGATTCAGTAATTTTCACCTCTCTAGTGTGGAAGCT 960  
Qy 961 TCACATAAGGATTTGTTCTGTTTCTCCATTCAGTGTGTTGTTATGTTGAGATCTTAAC 1020  
Db 961 TCACATAAGGATTTGTTCTGTTTCTCCATTCAGTGTGTTGTTATGTTGAGATCTTAAC 1020  
Qy 1021 CGTATCAATTTCTGTAATCAAACTTCTCTCTCTTTTGGAAAAAAGAAAAA 1079  
Db 1021 CGTATCAATTTCTGTAATGAAACTTCTCTCTCTTTTGGAAAAAAGAAAAA 1079

; Sequence 241778, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 241778  
; LENGTH: 607  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-241778

Query Match 2.7%; Score 29; DB 13; Length 607;  
Best Local Similarity 100.0%; Pred. No. 0.00041;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1051 TTCTTTTGGAAAAAAGAAAAA 1079  
Db 384 TTCTTTTGGAAAAAAGAAAAA 412

RESULT 3  
US-10-027-632-241778  
; Sequence 241778, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 241778  
; LENGTH: 607  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-241778

Query Match 2.7%; Score 29; DB 17; Length 607;  
Best Local Similarity 100.0%; Pred. No. 0.0041;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1051 TTCTTTTGAAGAAAAA 1079  
DB 384 TTCTTTTGAAGAAAAA 412

## RESULT 4

US-10-425-115-79873/c  
; Sequence 79873, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 79873  
; LENGTH: 615  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(615)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_17286C.1  
US-10-425-115-79873

Query Match 2.6%; Score 28; DB 20; Length 615;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1052 TCCTTTTGAAGAAAAA 1079  
DB 68 TCCTTTTGAAGAAAAA 41

## RESULT 5

US-10-242-535A-17538  
; Sequence 17538, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242,535A  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 10/085,783  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 17538  
; LENGTH: 207  
; TYPE: DNA  
; ORGANISM: Human  
US-10-242-535A-17538

Query Match 2.5%; Score 27; DB 17; Length 207;

Best Local Similarity 100.0%; Pred. No. 0.0044;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1053 CCTTTTGAAGAAAAA 1079  
DB 154 CCTTTTGAAGAAAAA 180

## RESULT 6

US-10-085-783A-17538  
; Sequence 17538, Application US/10085783A  
; Publication No. US20040037841A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 17538  
; LENGTH: 207  
; TYPE: DNA  
; ORGANISM: Human  
US-10-085-783A-17538

Query Match 2.5%; Score 27; DB 18; Length 207;  
Best Local Similarity 100.0%; Pred. No. 0.0044;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CCTTTTGAAGAAAAA 1079  
DB 154 CCTTTTGAAGAAAAA 180

## RESULT 7

US-09-983-965-4280/c  
; Sequence 4280, Application US/09983965  
; Patent No. US20020137160A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; FILE REFERENCE: 37-21(10297)C  
; CURRENT APPLICATION NUMBER: US/09/983,965  
; CURRENT FILING DATE: 2001-10-26  
; PRIOR APPLICATION NUMBER: US 09/465,231  
; PRIOR FILING DATE: 1998-12-15  
; PRIOR APPLICATION NUMBER: US 60/113,678  
; PRIOR FILING DATE: 1998-12-17  
; NUMBER OF SEQ ID NOS: 5912  
; SEQ ID NO 4280  
; LENGTH: 321  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 62-LIB3058-058-Q1-K1-H10  
US-09-983-965-4280

Query Match 2.5%; Score 27; DB 9; Length 321;  
Best Local Similarity 100.0%; Pred. No. 0.0045;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CCTTTTGAAGAAAAA 1079  
Db 47 CCTTTTGAAGAAAAA 21

## RESULT 8

US-09-873-367C-946/C  
; Sequence 946, Application US/09873367C  
; Publication No. US20030165839A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; APPLICANT: Soppet, Daniel  
; APPLICANT: Endress, Gregory  
; APPLICANT: Augustus, Meena  
; APPLICANT: Ebner, Reinhard  
; APPLICANT: Carter, Kenneth  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using  
; TITLE OF INVENTION: Signature Gene Sets  
; FILE REFERENCE: 689290-64  
; CURRENT APPLICATION NUMBER: US/09/873,367C  
; CURRENT FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: U.S. 60/236,891  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: U.S. 60/236,842  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: U.S. 60/244,867  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: U.S. 60/245,084  
; PRIOR FILING DATE: 2000-11-01  
; NUMBER OF SEQ ID NOS: 1067  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 946  
; LENGTH: 446  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-873-367C-946

Query Match 2.5%; Score 27; DB 10; Length 446;  
Best Local Similarity 100.0%; Pred. No. 0.0045;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CCTTTTGAAGAAAAA 1079  
Db 29 CCTTTTGAAGAAAAA 3

## RESULT 9

US-10-843-641A-946/C  
; Sequence 946, Application US/10843641A  
; Publication No. US2005006454A1  
; GENERAL INFORMATION:  
; APPLICANT: Avalon Pharmaceuticals, Inc.  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using  
; TITLE OF INVENTION: Signature Gene Sets  
; FILE REFERENCE: 689290-189  
; CURRENT APPLICATION NUMBER: US/10/843,641A  
; CURRENT FILING DATE: 2004-05-12  
; PRIOR APPLICATION NUMBER: US/09/873,367  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US/09/954,531  
; PRIOR FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US/09/954,456  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/09/962,436  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/09/962,832  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/09/964,824  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: US/09/967,768  
; PRIOR FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US/09/968,007  
; PRIOR FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: US/09/969,347  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: US/09/969,708  
; PRIOR FILING DATE: 2001-10-03  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 8447  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 946  
; LENGTH: 446  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-843-641A-946

Query Match 2.5%; Score 27; DB 21; Length 446;  
Best Local Similarity 100.0%; Pred. No. 0.0045;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CCTTTTGAAGAAAAA 1079  
Db 29 CCTTTTGAAGAAAAA 3

## RESULT 10

US-10-425-115-102591/C  
; Sequence 102591, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 102591  
; LENGTH: 511  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_25072C.1  
US-10-425-115-102591

Query Match 2.5%; Score 27; DB 20; Length 511;  
Best Local Similarity 100.0%; Pred. No. 0.0046;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CCTTTTGAAGAAAAA 1079  
Db 30 CCTTTTGAAGAAAAA 4

## RESULT 11

US-09-925-300-754  
; Sequence 754, Application US/09925300  
; Patent No. US20020151681A1  
; GENERAL INFORMATION:  
; APPLICANT: Craig Rosen,  
; APPLICANT: Steve Ruben  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA101  
; CURRENT APPLICATION NUMBER: US/09/925,300  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05988  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1890  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 754  
; LENGTH: 1795

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-300-754

Query Match          2.5%; Score 27; DB 9; Length 1795;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CCTTTTGAAGAAAAAAGAAAAA 1079
    |||||
DB 1757 CCTTTTGAAGAAAAAAGAAAAA 1783

RESULT 12
US-10-264-237-1253
; Sequence 1253, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn ver. 3.1
; SEQ ID NO 1253
; LENGTH: 1795
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-264-237-1253

Query Match          2.5%; Score 27; DB 17; Length 1795;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1053 CCTTTTGAAGAAAAAAGAAAAA 1079
    |||||
DB 1757 CCTTTTGAAGAAAAAAGAAAAA 1783

RESULT 13
US-10-242-535A-54828
; Sequence 54828, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 54828
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (20)..(20)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28)..(28)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-54828

Query Match          2.4%; Score 26; DB 18; Length 332;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 CTTTGTGAAAAAAGAAAAA 1079
    |||||
DB 304 CTTTGTGAAAAAAGAAAAA 329

RESULT 15
US-10-437-963-72087/c
; Sequence 72087, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28)..(28)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-54828

Query Match          2.4%; Score 26; DB 17; Length 332;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 CTTTGTGAAAAAAGAAAAA 1079
    |||||
DB 304 CTTTGTGAAAAAAGAAAAA 329

RESULT 14
US-10-085-783A-54828
; Sequence 54828, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 54828
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (20)..(20)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28)..(28)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-54828

Query Match          2.4%; Score 26; DB 18; Length 332;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 CTTTGTGAAAAAAGAAAAA 1079
    |||||
DB 304 CTTTGTGAAAAAAGAAAAA 329

RESULT 15
US-10-437-963-72087/c
; Sequence 72087, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
```

```
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 72087
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_72498C.1
US-10-437-963-72087

Query Match      2.4%; Score 26; DB 19; Length 379;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 CTTTGTGAAAAAAAAAAAAAAAAA 1079
      |||||||
Db 79 CTTTGTGAAAAAAAAAAAAAAAAA 54

Search completed: September 12, 2005, 20:49:57
Job time : 737.243 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 11:35:08 ; Search time 204.853 Seconds  
(without alignments)  
8618.590 Million cell updates/sec

Title: US-10-780-002-1  
Perfect score: 1079  
Sequence: 1 acgcgggaaagaaagaaaa.....gaaaaaaaaaaaaaaaaa 1079

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID                   | Description       |
|------------|-------|-------------|--------|----------------------|-------------------|
| 1          | 26    | 2.4         | 814    | US-09-780-717-34     | Sequence 34, Appl |
| 2          | 25    | 2.3         | 235    | US-09-621-976-9455   | Sequence 9455, Ap |
| 3          | 25    | 2.3         | 492    | US-09-270-767-6495   | Sequence 6495, Ap |
| 4          | 25    | 2.3         | 492    | US-09-270-767-21777  | Sequence 21777, A |
| 5          | 25    | 2.3         | 1096   | US-09-136-073-1      | Sequence 1, Appli |
| 6          | 25    | 2.3         | 1096   | US-09-457-024A-1     | Sequence 1, Appli |
| 7          | 25    | 2.3         | 1811   | US-08-808-931-9      | Sequence 9, Appli |
| 8          | 25    | 2.3         | 1811   | US-08-808-323-9      | Sequence 9, Appli |
| 9          | 25    | 2.3         | 1811   | US-09-050-603A-9     | Sequence 9, Appli |
| 10         | 25    | 2.3         | 1811   | US-09-102-420B-9     | Sequence 9, Appli |
| 11         | 25    | 2.3         | 1811   | US-09-497-698-9      | Sequence 9, Appli |
| 12         | 25    | 2.3         | 1811   | US-09-730-525-9      | Sequence 9, Appli |
| 13         | 25    | 2.3         | 2080   | US-09-311-021-179    | Sequence 179, App |
| 14         | 25    | 2.3         | 3138   | US-09-234-332-5      | Sequence 5, Appli |
| 15         | 25    | 2.3         | 29930  | US-09-949-016-15326  | Sequence 15326, A |
| 16         | 25    | 2.3         | 636591 | US-09-949-016-11808  | Sequence 11808, A |
| 17         | 25    | 2.3         | 636591 | US-09-949-016-13388  | Sequence 13388, A |
| 18         | 24    | 2.2         | 282    | US-09-621-976-18648  | Sequence 18648, A |
| 19         | 24    | 2.2         | 299    | US-09-621-976-7775   | Sequence 7775, Ap |
| 20         | 24    | 2.2         | 394    | US-09-621-976-15204  | Sequence 15204, A |
| 21         | 24    | 2.2         | 517    | US-10-101-464A-970   | Sequence 970, App |
| 22         | 24    | 2.2         | 601    | US-09-949-016-113460 | Sequence 113460   |
| 23         | 24    | 2.2         | 1039   | US-09-464-535-23     | Sequence 23, Appl |
| 24         | 24    | 2.2         | 1140   | US-08-698-805-5      | Sequence 5, Appli |
| 25         | 24    | 2.2         | 1193   | US-09-347-798-1      | Sequence 1, Appli |
| 26         | 24    | 2.2         | 1699   | US-09-152-060-19     | Sequence 19, Appl |
| 27         | 24    | 2.2         | 1745   | US-08-464-523B-1     | Sequence 1, Appli |

|    |    |     |       |   |                     |                   |
|----|----|-----|-------|---|---------------------|-------------------|
| 28 | 24 | 2.2 | 2082  | 3 | US-09-149-476-257   | Sequence 257, App |
| 29 | 24 | 2.2 | 2434  | 4 | US-09-820-007-1     | Sequence 1, Appli |
| 30 | 24 | 2.2 | 3183  | 4 | US-09-220-132-16    | Sequence 16, Appl |
| 31 | 24 | 2.2 | 3722  | 4 | US-10-164-595-9     | Sequence 9, Appli |
| 32 | 24 | 2.2 | 3855  | 3 | US-08-974-549A-4    | Sequence 4, Appli |
| 33 | 24 | 2.2 | 3855  | 4 | US-08-912-951-4     | Sequence 4, Appli |
| 34 | 24 | 2.2 | 3855  | 4 | US-09-402-181B-4    | Sequence 4, Appli |
| 35 | 24 | 2.2 | 3855  | 4 | US-09-721-456-4     | Sequence 4, Appli |
| 36 | 24 | 2.2 | 3862  | 4 | US-10-164-595-5     | Sequence 5, Appli |
| 37 | 24 | 2.2 | 3937  | 4 | US-10-164-595-7     | Sequence 7, Appli |
| 38 | 24 | 2.2 | 3985  | 4 | US-10-164-595-3     | Sequence 3, Appli |
| 39 | 24 | 2.2 | 4029  | 3 | US-08-851-843A-173  | Sequence 173, App |
| 40 | 24 | 2.2 | 4029  | 3 | US-08-974-549A-292  | Sequence 292, App |
| 41 | 24 | 2.2 | 4029  | 3 | US-08-854-050-173   | Sequence 173, App |
| 42 | 24 | 2.2 | 4029  | 3 | US-09-430-323-173   | Sequence 173, App |
| 43 | 24 | 2.2 | 4029  | 4 | US-09-402-181B-292  | Sequence 292, App |
| 44 | 24 | 2.2 | 4029  | 4 | US-09-721-456-292   | Sequence 292, App |
| 45 | 24 | 2.2 | 30221 | 4 | US-09-949-016-12577 | Sequence 12577, A |

ALIGNMENTS

RESULT 1  
US-09-780-717-34  
; Sequence 34, Application US/09780717  
; Patent No. 6713666  
; GENERAL INFORMATION:  
; APPLICANT: Heientjaris, Tim  
; APPLICANT: Bates, Nic  
; APPLICANT: Allen, Stephen M.  
; TITLE OF INVENTION: No. 6713666el Invertase Inhibitors and Methods  
; TITLE OF INVENTION: of Use  
; FILE REFERENCE: 035718/208677  
; CURRENT APPLICATION NUMBER: US/09/780,717  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: US 60/181,509  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 34  
; LENGTH: 814  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (99)...(638)  
US-09-780-717-34

Query Match 2.4%; Score 26; DB 4; Length 814;  
Best Local Similarity 100.0%; Pred. No. 0.15;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1054 CTTTGTGAAAAAAAAAAAAAAAAAAAA 1079  
|||||  
Db 770 CTTTGTGAAAAAAAAAAAAAAAAAAAA 795

RESULT 2  
US-09-621-976-9455  
; Sequence 9455, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 9455

```
; LENGTH: 235
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-9455

Query Match          2.3%; Score 25; DB 4; Length 235;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 TTTTGGAAAAA 1079
Db 176 TTTTGGAAAAA 200

RESULT 3
US-09-270-767-6495
; Sequence 6495, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6495
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-6495

Query Match          2.3%; Score 25; DB 4; Length 492;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 TTTTGGAAAAA 1079
Db 385 TTTTGGAAAAA 409

RESULT 4
US-09-270-767-21777
; Sequence 21777, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21777
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-21777

Query Match          2.3%; Score 25; DB 4; Length 492;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 TTTTGGAAAAA 1079
Db 385 TTTTGGAAAAA 409

RESULT 5
US-09-136-073-1
; Sequence 1, Application US/09136073
; Patent No. 6043076
; GENERAL INFORMATION:
```

```
; APPLICANT: Ramakrishnan, Santha
; TITLE OF INVENTION: GENE ENCODING 2,3-DIHYDROXYBENZOIC ACID DECARBOXYLASE
; FILE REFERENCE: UNVN-53687
; CURRENT APPLICATION NUMBER: US/09/136,073
; CURRENT FILING DATE: 1998-08-18
; EARLIER APPLICATION NUMBER: 60/056,621
; EARLIER FILING DATE: 1997-08-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Corel WordPerfect 6.1 for Windows
; SEQ ID NO 1
; LENGTH: 1096
; TYPE: DNA
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: gene
; LOCATION: 1-1096
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1-1029
; FEATURE:
; NAME/KEY: polyA site
; LOCATION: 1073
; FEATURE:
; NAME/KEY: polyA signal
; LOCATION: 1056-1061
; FEATURE:
; NAME/KEY: 3' UTR
; LOCATION: 1027-1073
; FEATURE:
; NAME/KEY: source/Aspergillus niger
; LOCATION: 1-1096
; PUBLICATION INFORMATION:
; AUTHORS: Santha, Ramakrishnan
; AUTHORS: Dickman, Martin B.
; AUTHORS: O'Leary, Marion H.
; TITLE: 2,3-Dihydroxybenzoic Acid Decarboxylase From Aspergillus niger:
; TITLE: Mechanism, Cloning And Overexpression.
; JOURNAL: Faseb Journal
; VOLUME: 11
; ISSUE: 9
; PAGES: A1017
; DATE: 1997-07-31
US-09-136-073-1

Query Match          2.3%; Score 25; DB 3; Length 1096;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 TTTTGGAAAAA 1079
Db 1071 TTTTGGAAAAA 1095

RESULT 6
US-09-457-024A-1
; Sequence 1, Application US/09457024A
; Patent No. 6440704
; GENERAL INFORMATION:
; APPLICANT: Ramakrishnan, Santha
; TITLE OF INVENTION: GENE ENCODING 2,3-DIHYDROXYBENZOIC ACID DECARBOXYLASE
; FILE REFERENCE: UNVN-53687
; CURRENT APPLICATION NUMBER: US/09/457,024A
; CURRENT FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: 09/136,073
; PRIOR FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Corel WordPerfect 6.1 for Windows
; SEQ ID NO 1
; LENGTH: 1096
; TYPE: DNA
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: gene
```



```
; LOCATION: 1-1096
; NAME/KEY: CDS
; LOCATION: 1-1029
; NAME/KEY: polyA_site
; LOCATION: 1073
; NAME/KEY: polyA_signal
; LOCATION: 1056-1061
; NAME/KEY: 3' UTR
; LOCATION: 1027-1073
; NAME/KEY: source/Aspergillus niger
; LOCATION: 1-1096
; PUBLICATION INFORMATION:
; AUTHORS: Santha, Ramakrishnan
; AUTHORS: Dickman, Martin B.
; AUTHORS: O'Leary, Marion H.
; TITLE: 2,3-Dihydroxybenzoic Acid Decarboxylase From Aspergillus niger:
; TITLE: Mechanism, Cloning And Overexpression.
; JOURNAL: Faseb Journal
; VOLUME: 11
; ISSUE: 9
; PAGES: A1017
; DATE: 1997-07-31
; US-09-457-024A-1
```

```
Query Match 2.3%; Score 25; DB 3; Length 1096;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1055 TTTTGGAAAAA 1079
Db 1071 TTTTGGAAAAA 1095
```

```
RESULT 7
US-08-808-931-9
; Sequence 9, Application US/08808931
; Patent No. 5939602
; GENERAL INFORMATION:
; APPLICANT: Volrath, Sandra
; APPLICANT: Johnson, Marie
; APPLICANT: Potter, Sharon
; APPLICANT: Ward, Eric
; APPLICANT: Helfetz, Peter
; TITLE OF INVENTION: DNA Molecules Encoding Plant
; TITLE OF INVENTION: Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,931
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,705
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,612
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1846
```

```
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1847
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1811 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Triticum aestivum (wheat)
; IMMEDIATE SOURCE:
; CLONE: pWDC-13 (NRRL B-21545)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1589
; OTHER INFORMATION: /product= "wheat protox-1"
; US-08-808-931-9
```

```
Query Match 2.3%; Score 25; DB 2; Length 1811;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1055 TTTTGGAAAAA 1079
Db 1787 TTTTGGAAAAA 1811
```

```
RESULT 8
US-08-808-323-9
; Sequence 9, Application US/08808323
; Patent No. 6018105
; GENERAL INFORMATION:
; APPLICANT: Johnson, Marie
; APPLICANT: Volrath, Sandra
; APPLICANT: Ward, Eric
; TITLE OF INVENTION: Promoters from Plant
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 6018105artis Corporation
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,323
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,705
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,612
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1846
```

```

/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (919) 541-8587
/ TELEFAX: (919) 541-8689
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1811 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ HYPOTHETICAL: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Triticum aestivum (wheat)
/ IMMEDIATE SOURCE:
/ CLONE: pWDC-13 (NRRL B-21545)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 3..1589
/ OTHER INFORMATION: /product= "wheat protox-1"
/
US-08-808-323-9

Query Match 2.3%; Score 25; DB 3; Length 1811;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1055 TTTTGAAGAAAAA 1079
Db 1787 TTTTGAAGAAAAA 1811

RESULT 9
US-09-050-603A-9
/ Sequence 9, Application US/09050603A
/ Patent No. 6023012
/ GENERAL INFORMATION:
/ APPLICANT: Volrath, Sandra
/ APPLICANT: Johnson, Marie
/ APPLICANT: Potter, Sharon
/ APPLICANT: Ward, Eric
/ APPLICANT: Heifetz, Peter
/ TITLE OF INVENTION: DNA Molecules Encoding Plant
/ TITLE OF INVENTION: Protoporphyrogen Oxidase
/ NUMBER OF SEQUENCES: 37
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: No. 6023012artis Corporation
/ STREET: 3054 Cornwallis Road
/ CITY: Research Triangle Park
/ STATE: NC
/ COUNTRY: USA
/ ZIP: 27709
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION NUMBER: US/09/050,603A
/ FILING DATE: 30-MAR-1998
/ CLASSIFICATION: 800
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/808,931
/ FILING DATE: 28-FEB-1997
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 60/012,705
/ FILING DATE: 28-FEB-1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 60/013,612
/ FILING DATE: 28-FEB-1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 60/020,003
/ FILING DATE: 21-JUN-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Meigs, J. Timothy
/
US-09-050-603A-9
/ Sequence 9, Application US/09102420B
/ Patent No. 6084155
/ GENERAL INFORMATION:
/ APPLICANT: Volrath, Sandra
/ APPLICANT: Johnson, Marie
/ APPLICANT: Ward, Eric
/ APPLICANT: Heifetz, Peter
/ TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
/ TITLE OF INVENTION: OXIDASE ("PROTOX")
/ NUMBER OF SEQUENCES: 43
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: No. 6084155artis Corporation
/ STREET: 3054 Cornwallis Road
/ CITY: Research Triangle Park
/ STATE: NC
/ COUNTRY: USA
/ ZIP: 27709
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION NUMBER: US/09/102,420B
/ FILING DATE: 22-JUN-1998
/ CLASSIFICATION: 800
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 09/059,164
/ FILING DATE: 13-APR-1998
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 09/050,603
/ FILING DATE: 30-MAR-1998
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 60/126,430
/ FILING DATE: 11-MAR-1998
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/808,931
/ FILING DATE: 28-FEB-1997
/ PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: US 60/012,705  
FILING DATE: 28-FEB-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/013,612  
FILING DATE: 28-FEB-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/020,003  
FILING DATE: 21-JUN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/472,028  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: CGC 1847/CIP3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919) 541-8587  
TELEFAX: (919) 541-8689  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1811 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Triticum aestivum (wheat)  
IMMEDIATE SOURCE:  
CLONE: pWDC-13 (NRRL B-21545)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3..1589  
OTHER INFORMATION: /product= "wheat protox-1"  
US-09-102-420B-9

Query Match 2.3%; Score 25; DB 3; Length 1811;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 TTTTGAAGAAAAAAGAAAAA 1079  
|||||  
DB 1787 TTTTGAAGAAAAAAGAAAAA 1811  
|||||

RESULT 11  
US-09-497-698-9  
Sequence 9, Application US/09497698  
Patent No. 6308458  
GENERAL INFORMATION:  
APPLICANT: Volrath, Sandra  
Johnson, Marie  
Ward, Eric  
Heifetz, Peter  
TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN  
OXIDASE ("PROTOX")  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6308458artis Corporation  
STREET: 3054 Cornwalis Road  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/497,698  
FILING DATE: 03-Feb-2000  
CLASSIFICATION: <Unknown>

30-MAR-1998  
11-MAR-1998  
28-FEB-1997  
28-FEB-1996  
28-FEB-1996  
21-JUN-1996  
06-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/102,420  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 09/050,603  
FILING DATE: 30-MAR-1998  
APPLICATION NUMBER: US 60/126,430  
FILING DATE: 11-MAR-1998  
APPLICATION NUMBER: US 08/808,931  
FILING DATE: 28-FEB-1997  
APPLICATION NUMBER: US 60/012,705  
FILING DATE: 28-FEB-1996  
APPLICATION NUMBER: US 60/013,612  
FILING DATE: 28-FEB-1996  
APPLICATION NUMBER: US 60/020,003  
FILING DATE: 21-JUN-1996  
APPLICATION NUMBER: US 08/472,028  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: CGC 1847/CIP3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919) 541-8587  
TELEFAX: (919) 541-8689  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1811 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Triticum aestivum (wheat)  
IMMEDIATE SOURCE:  
CLONE: pWDC-13 (NRRL B-21545)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3..1589  
OTHER INFORMATION: /product= "wheat protox-1"  
US-09-497-698-9

Query Match 2.3%; Score 25; DB 3; Length 1811;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 TTTTGAAGAAAAAAGAAAAA 1079  
|||||  
DB 1787 TTTTGAAGAAAAAAGAAAAA 1811  
|||||

RESULT 12  
US-09-730-525-9  
Sequence 9, Application US/09730525  
Patent No. 6808904  
GENERAL INFORMATION:  
APPLICANT: Ward, Eric  
Volrath, Sandra  
Johnson, Marie  
Potter, Sharon  
TITLE OF INVENTION: Herbicide Tolerant Protox Genes  
Produced by DNA Shuffling  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6808904artis Corporation

```
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/730,525
FILING DATE: 05-Dec-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/102,419
FILING DATE: <Unknown>
APPLICATION NUMBER: US 09/050,603
FILING DATE: 30-MAR-1998
APPLICATION NUMBER: US 09/038,878
FILING DATE: 11-MAR-1998
APPLICATION NUMBER: US 08/808,931
FILING DATE: 28-FEB-1997
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
APPLICATION NUMBER: US 08/472,028
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847/CIP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1811 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Triticum aestivum (wheat)
IMMEDIATE SOURCE:
CLONE: pWDC-13 (NRRL B-21545)
FEATURE:
NAME/KEY: CDS
LOCATION: 3..1589
OTHER INFORMATION: /product= "wheat protox-1"
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-730-525-9
Query Match 2.3%; Score 25; DB 4; Length 1811;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1055 TTTTGAAGAAAAA 1079
DB 1787 TTTTGAAGAAAAA 1811
RESULT 13
US-09-311-021-179
; Sequence 179, Application US/09311021
; Patent No. 6706869
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
```

```
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6300-11A
; CURRENT APPLICATION NUMBER: US/09/311,021
; CURRENT FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 179
; LENGTH: 2080
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-311-021-179
Query Match 2.3%; Score 25; DB 4; Length 2080;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1055 TTTTGAAGAAAAA 1079
DB 2045 TTTTGAAGAAAAA 2069
RESULT 14
US-09-234-332-5/c
; Sequence 5, Application US/09234332A
; Patent No. 6087168
; GENERAL INFORMATION:
; APPLICANT: Cedars-Sinai Medical Center
; APPLICANT: Michel F. Levesque, M.D.
; APPLICANT: Thomas Neuman, Ph.D.
; TITLE OF INVENTION: CONVERSION OF NON-NEURONAL CELLS INTO
; TITLE OF INVENTION: NEURONS; TRANSDIFFERENTIATION OF EPIDERMAL CELLS
; FILE REFERENCE: P07 41494
; CURRENT APPLICATION NUMBER: US/09/234,332A
; CURRENT FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 3138
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: gene
; LOCATION: (0)...(0)
; OTHER INFORMATION: Zic 1 Protein gene; Genbank Acession D76435
US-09-234-332-5
Query Match 2.3%; Score 25; DB 3; Length 3138;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1054 CTTTGTGAAAAA 1078
DB 219 CTTTGTGAAAAA 195
RESULT 15
US-09-949-016-15326/c
; Sequence 15326, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
```



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: September 12, 2005, 08:41:37 ; Search time 3440.77 Seconds  
(without alignments)  
11936.677 Million cell updates/sec

Title: US-10-780-002-1  
Perfect score: 1079  
Sequence: 1 acgcgggaagaagaagaa.....gaaaaaaaaaaaaaaaaa 1079

Scoring table: OLIGO\_NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Match | Length | DB | ID       | Description |
|------------|-------|-------|--------|----|----------|-------------|
| 1          | 49    | 4-5   | 494    | 2  | BE924354 | EST428123   |
| 2          | 41    | 3-8   | 384    | 2  | BF097870 | EST415943   |
| 3          | 41    | 3-8   | 646    | 2  | AW621893 | EST312691   |
| 4          | 34    | 3-2   | 420    | 1  | AI772824 | EST253924   |
| 5          | 34    | 3-2   | 476    | 1  | AI772823 | EST253923   |
| 6          | 34    | 3-2   | 535    | 2  | BE433359 | EST399888   |
| 7          | 33    | 3-1   | 922    | 7  | CK269872 | EST715950   |
| 8          | 31    | 2-9   | 525    | 1  | AI775631 | EST256711   |
| 9          | 31    | 2-9   | 546    | 2  | AW224223 | EST301130   |
| 10         | 31    | 2-9   | 590    | 4  | BI928709 | EST548598   |
| 11         | 31    | 2-9   | 781    | 4  | BG126704 | EST472350   |
| 12         | 30    | 2-8   | 124    | 2  | BB583163 | BB583163    |
| 13         | 30    | 2-8   | 388    | 1  | AI591310 | tw11f05.x   |
| 14         | 30    | 2-8   | 400    | 1  | AI718513 | as56h02.x   |
| 15         | 30    | 2-8   | 417    | 5  | BQ383556 | NISC.mm01   |
| 16         | 30    | 2-8   | 419    | 5  | BQ159741 | WHE2228.C   |
| 17         | 30    | 2-8   | 472    | 5  | BY240192 | BY240192    |
| 18         | 30    | 2-8   | 535    | 4  | BG804653 | 0264-53.M   |
| 19         | 30    | 2-8   | 600    | 4  | BG806457 | 0264-53.M   |
| 20         | 30    | 2-8   | 600    | 5  | BU924647 | 7094-43.M   |
| 21         | 30    | 2-8   | 657    | 1  | AL897127 | AL897127    |
| 22         | 30    | 2-8   | 684    | 6  | CA316959 | UI-M-FW0-   |
| 23         | 30    | 2-8   | 698    | 7  | CN460297 | UI-M-HB0-   |
| 24         | 30    | 2-8   | 855    | 5  | BX683276 | BX683276    |

```

c 25 30 2.8 938 5 BUS20408
c 26 29 2.7 207 6 C93102
c 27 29 2.7 364 1 A1431179
c 28 29 2.7 399 7 CV496924
c 29 2.7 548 6 CA814690
c 30 29 2.7 614 7 CV258573
c 31 29 2.7 966 7 CN644624
c 32 28 2.6 218 2 BB589554
c 33 28 2.6 227 7 COS11445
c 34 28 2.6 294 2 BF651915
c 35 28 2.6 434 4 BM392172
c 36 28 2.6 448 4 BI495836
c 37 28 2.6 468 1 A1681388
c 38 28 2.6 511 7 CK433247
c 39 28 2.6 535 1 A1765652
c 40 28 2.6 550 7 CK433514
c 41 28 2.6 561 7 CV225559
c 42 28 2.6 581 7 CK433735
c 43 28 2.6 622 7 CR585931
c 44 28 2.6 627 6 CA423834
c 45 28 2.6 694 4 BI294974

```

## ALIGNMENTS

```

RESULT 1
BE924354
LOCUS EST428123 potato leaves and petioles Solanum tuberosum cDNA clone
DEFINITION CSTB29K16 5' sequence, mRNA sequence.
ACCESSION BE924354
VERSION BE924354.1 GI:10450430
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
REFERENCE 1 (bases 1 to 494)
AUTHORS van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R.,
Ronnig,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
TITLE Generation of ESTs from potato leaves and petioles
JOURNAL Unpublished (2000)
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/.

```

```

FEATURES
source
location/Qualifiers
1..494
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTB29K16"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/clone_lib="potato leaves and petioles"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."

```

## ORIGIN

Query Match 4.5%; Score 49; DB 2; Length 494;  
Best Local Similarity 100.0%; Pred. No. 8.9e-13;

```

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 271 TAGTGGGGCATAGTCTTGTTGGTATGAATTTGGGACTTGCTATGAAAA 319
|||||
Db 258 TAGTGGGGCATAGTCTTGTTGGTATGAATTTGGGACTTGCTATGAAAA 306
|||||

RESULT 2
BF097870 384 bp mRNA linear EST 18-MAY-2001
DEFINITION EST415943 tomato nutrient deficient roots Lycopersicon esculentum
CDNA clone CLEW24G19 5' sequence, mRNA sequence.
BF097870
EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 384)
REFERENCE van der Hoeven,R.S., Garvin,D.P., Matern,A.L., Holt,I.E., Liang,F.,
AUTHORS Upton,J., Hansen,T.S., Ronning,C.M., Craven,M.B., Bowman,C.L.,
Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Giovannoni,J.J.
and Tanksley,S.D.
Generation of ESTs from tomato nutrient-deficient roots
Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
Location/Qualifiers
1..384
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA492"
/db_xref="taxon:4081"
/clone="CLEW24G19"
/tissue_type="roots"
/dev_stage="5-6 weeks old"
/lab_host="SOL8"
/clone_lib="tomato nutrient deficient roots"
/note="Vector: pBluescriptSMCudapt; Site 1: 5' EcoRI;
Site 2: 3' XhoI; Roots were harvested from plants grown
under the following deficiencies/stresses: 10 mM Al, Zn,
P, K. Fe.N. mRNA was isolated from individual treatments.
Proportional aliquots of mRNA of each treatment were mixed
and used for library construction."

ORIGIN
Query Match 3.8%; Score 41; DB 2; Length 384;
Best Local Similarity 100.0%; Pred. No. 8.6e-09;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 CATAGTCTTGGTATGAATTTGGGACTTGCTATGAAAA 319
|||||
Db 249 CATAGTCTTGGTATGAATTTGGGACTTGCTATGAAAA 289
|||||

RESULT 3
AW621893 646 bp mRNA linear EST 18-MAY-2001
LOCUS EST312691 tomato root during/after fruit set, Cornell University
DEFINITION Lycopersicon esculentum cDNA clone CLEX13L15 5', mRNA sequence.
AW621893
ACCESSION AW621893
VERSION AW621893
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

```

```

REFERENCE 1 (bases 1 to 646)
AUTHORS van der Hoeven,R.S., Garvin,D., Matern,A.L., Holt,I.E., Kochian,L.,
Liang,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and
Tanksley,S.D.
Generation of ESTs from tomato root, during and after fruit set
Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
Location/Qualifiers
1..646
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CLEX13L15"
/tissue_type="root"
/dev_stage="plants during and after fruit-set"
/clone_lib="tomato root during/after fruit set, Cornell
University"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: tanksley; tissue supplied by Dave Garvin
(USDA-ARS, Ithaca, NY 14850)."

ORIGIN
Query Match 3.8%; Score 41; DB 2; Length 646;
Best Local Similarity 100.0%; Pred. No. 8.5e-09;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 CATAGTCTTGGTATGAATTTGGGACTTGCTATGAAAA 319
|||||
Db 249 CATAGTCTTGGTATGAATTTGGGACTTGCTATGAAAA 289
|||||

RESULT 4
AW1772824 420 bp mRNA linear EST 18-MAY-2001
LOCUS EST253924 tomato resistant, Cornell Lycopersicon esculentum cDNA
DEFINITION clone CLER4A4, mRNA sequence.
AW1772824
ACCESSION AW1772824 GI:5270865
VERSION AW1772824.1
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 420)
REFERENCE D'Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,
AUTHORS Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y.,
Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,
Giovannoni,J.J. and Martin,G.B.
Generation of ESTs from Pseudomonas resistant tomato
Unpublished (1999)
Other ESTs: TC2068
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
Location/Qualifiers
1..420
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="R11-12 (35S::Pto in Rio Grande x Money Maker)"
/db_xref="taxon:4081"
/clone="CLER4A4"
/tissue_type="leaf"

```



```

/dev_stage="4-week old"
/lab_host="SOLR"
/clone_lib="tomato resistant, Cornell"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLSP - Tomato Pseudomonas Resistant EST Library.
Directionally cloned cDNAs inserted into pBluescript
SK(-) at 5' end with EcoRI and 3' end with XhoI site."

```

## ORIGIN

```

Query Match      3.2%; Score 34; DB 1; Length 420;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 328 AAAAGATCTATGCTGCTGTTTCTTGGCTGCTTT 361

Db 302 AAAAGATCTATGCTGCTGTTTCTTGGCTGCTTT 335

## RESULT 5

AI772823

```

LOCUS      AI772823      476 bp      mRNA      linear      EST 18-MAY-2001
DEFINITION EST253923 tomato resistant, Cornell Lycopersicon esculentum cDNA
ACCESSION  clone cLER4A2, mRNA sequence.
VERSION    AI772823      1 GI:5270864
KEYWORDS
SOURCE     Lycopersicon esculentum (tomato)
ORGANISM

```

EuKaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 476)

REFERENCE

AUTHORS

D'Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E., Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Niemman, W., Fraser, C.M., Venter, J.C., Tanksley, S.D., Giovannoni, J.J. and Martin, G.B.

Generation of ESTs from Pseudomonas resistant tomato

Unpublished (1999)

Other ESTs: TC2068

Contact: CUGI

Clemson University

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

## FEATURES

source

```

Location/Qualifiers
1..476
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="R11-12 (35S::Pto in Rio Grande x Money Maker)"
/db_xref="taxon:4081"
/clone="cLER4A2"
/tissue_type="leaf"
/dev_stage="4-week old"
/lab_host="SOLR"
/clone_lib="tomato resistant, Cornell"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLSP - Tomato Pseudomonas Resistant EST Library.
Directionally cloned cDNAs inserted into pBluescript
SK(-) at 5' end with EcoRI and 3' end with XhoI site."

```

## ORIGIN

```

Query Match      3.2%; Score 34; DB 1; Length 476;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 328 AAAAGATCTATGCTGCTGTTTCTTGGCTGCTTT 361

Db 301 AAAAGATCTATGCTGCTGTTTCTTGGCTGCTTT 334

## RESULT 6

BE433359

```

LOCUS      BE433359      535 bp      mRNA      linear      EST 18-MAY-2001
DEFINITION EST399888 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
ACCESSION  clone cLEG13E6, mRNA sequence.
VERSION    BE433359      1 GI:9431202
KEYWORDS
SOURCE     Lycopersicon esculentum (tomato)
ORGANISM

```

EuKaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 535)

REFERENCE

AUTHORS

Alcala, J., Vrebalov, J., White, R., van der Hoeven, R.S., Holt, I.E., Liang, F., Hansen, T.S., Craven, M.B., Bowman, C.L., Ronning, C.M., Niemman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.

Generation of ESTs from tomato fruit tissue, breaker stage

Unpublished (2000)

Contact: CUGI

Clemson University

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

## FEATURES

source

```

Location/Qualifiers
1..535
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEG13E6"
/tissue_type="pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/clone_lib="tomato breaker fruit, TIGR"
/notes="Vector: pBluescriptSMCquadapt; Site 1: EcoRI;
Site 2: XhoI; Fruit were harvested at the breaker stage
(first sign of lycopen accumulation on the blossom end of
the fruit). Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."

```

## ORIGIN

```

Query Match      3.2%; Score 34; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 328 AAAAGATCTATGCTGCTGTTTCTTGGCTGCTTT 361

Db 298 AAAAGATCTATGCTGCTGTTTCTTGGCTGCTTT 331

## RESULT 7

CK269872

LOCUS

```

DEFINITION EST715950 potato abiotic stress cDNA library Solanum tuberosum cDNA
ACCESSION  clone POACP55 5' end, mRNA sequence.
VERSION    CK269872      1 GI:39826850
KEYWORDS
SOURCE     Solanum tuberosum (potato)
ORGANISM

```

EuKaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.

1 (bases 1 to 922)

REFERENCE

AUTHORS

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.

Generation of ESTs from abiotic stressed potato tissue

Unpublished (2003)

Other ESTs: EST715951

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: [potato-array@tigr.org](mailto:potato-array@tigr.org)

Clones can be requested from the University of Arizona Genomics Institute via <http://genome.arizona.edu/orders/>.  
Seq primer: ATT TAG GTG ACA CTA TAG.

#### FEATURES source

Location/Qualifiers  
1. .922  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POACP55"  
/tissue\_type="abiotic stress treated leaf and root tissue"  
/lab\_host="DH10B-TonA"

/clone\_lib="potato abiotic stress cDNA library"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

#### ORIGIN

Query Match 3.1%; Score 33; DB 7; Length 922;  
Best Local Similarity 100.0%; Pred. No. 8.1e-05;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 287 TGGTGTATGAATTTGGGACTTGGCTATGGAAAA 319  
|||||  
Db 279 TGGTGTATGAATTTGGGACTTGGCTATGGAAAA 311

#### RESULT 8 AI775631

LOCUS  
DEFINITION  
AI775631 525 bp mRNA linear EST 18-MAY-2001  
clone cLER1615, mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Lycopersicon esculentum (tomato)  
Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE  
AUTHORS  
1 (bases 1 to 525)  
D'Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E.,  
Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y.,  
Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Tanksley, S.D.,  
Giovannoni, J.J., and Martin, G.B.

TITLE  
JOURNAL  
COMMENT  
Generation of ESTs from Pseudomonas resistant tomato  
Contact: CUGI  
Unpublished (1999)

Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.

#### FEATURES source

Location/Qualifiers  
1. .525  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="R11-12 (35S)::Pto in Rio Grande x Money Maker)"  
/db\_xref="taxon:4081"  
/clone="cLER1615"  
/tissue\_type="leaf"  
/dev\_stage="4-week old"  
/lab\_host="SOLR"  
/clone\_lib="tomato resistant, Cornell"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
XhoI; cLER - Tomato Pseudomonas Resistant EST library.  
Directionally cloned cDNAs inserted into pBluescript  
SK(-) at 5' end with EcoRI and 3' end with XhoI site."

#### ORIGIN

Query Match 2.9%; Score 31; DB 1; Length 525;  
Best Local Similarity 100.0%; Pred. No. 0.00082;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 279 CATAGTCTTGGTGTATGAATTTGGACTTG 309  
|||||  
Db 274 CATAGTCTTGGTGTATGAATTTGGACTTG 304

#### RESULT 9 AW224223

LOCUS  
DEFINITION  
AW224223 546 bp mRNA linear EST 18-MAY-2001  
clone cLEN16B20, mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Lycopersicon esculentum (tomato)  
Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE  
AUTHORS  
1 (bases 1 to 546)  
Alcala, J., Vrebalov, J., White, R., Matern, A.L., Holt, I.E., Liang, F.,  
Upton, J., Hansen, T., Craven, M.B., Bowman, C.L., Ahn, S.,  
Ronning, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and  
Giovannoni, J.  
Generation of ESTs from tomato fruit tissue  
Unpublished (1999)  
Contact: CUGI

Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.

#### FEATURES source

Location/Qualifiers  
1. .546  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cLEN16B20"  
/tissue\_type="pericarp"  
/dev\_stage="red ripe (7-20 days post-breaker)"  
/clone\_lib="tomato fruit red ripe, TAMU"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
XhoI; supplier: Giovannoni; Fruit were tagged at the  
breaker stage (first sign of lycopene accumulation on the  
blossom end of the fruit) and harvested 7 days  
post-breaker (fully red-ripe), 10 days post breaker, and  
20 days post-breaker (over-ripe). 20 day fruit which  
showed external or internal signs of pathogenesis were  
discarded. Fruit were cut in half and the seeds and  
locules were discarded prior to freezing the pericarp."

#### ORIGIN

```

Query Match      2.9%; Score 31; DB 2; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.00082;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 CATAGCTTGGTGTATGAATTTGGGACTTG 309
|||||
Db 358 CATAGCTTGGTGTATGAATTTGGGACTTG 388
|||||

RESULT 10
BI928709
LOCUS EST548598 tomato flower, 3 - 8 mm buds Lycopersicon esculentum cDNA
DEFINITION clone cTOB26K14 5' end, mRNA sequence.
ACCESSION BI928709
VERSION BI928709.1 GI:16240757
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 590)
AUTHORS van der Hoeven,R.S., Bezzerides,J.L., Karanycheva,S.A., Tsai,J.,
Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
TITLE Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.
Location/Qualifiers
1..590
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOB26K14"
/tissue_type="flower"
/dev_stage="3-8mm buds"
/clone_lib="tomato flower, 3 - 8 mm buds"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research. Flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

ORIGIN
Query Match      2.9%; Score 31; DB 4; Length 590;
Best Local Similarity 100.0%; Pred. No. 0.00082;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 CATAGCTTGGTGTATGAATTTGGGACTTG 309
|||||
Db 272 CATAGCTTGGTGTATGAATTTGGGACTTG 302
|||||

RESULT 11
BG126704
LOCUS EST472350 tomato shoot/meristem Lycopersicon esculentum cDNA clone
DEFINITION cTOF13K13 5' sequence, mRNA sequence.
ACCESSION BG126704
VERSION BG126704.1 GI:12626892
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 781)
AUTHORS van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Utterback,T.,
Hansen,C., Ronning,C. and Tanksley,S.
TITLE Generation of ESTs from tomato shoot/meristem tissue
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
Location/Qualifiers
1..781
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOF13K13"
/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/lab_host="SOLR"
/clone_lib="tomato shoot/meristem"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Small expanding leaves from the growing tip were
taken from greenhouse plants (4-6wks old TA496). Tissue
was immediately frozen in liquid nitrogen."

ORIGIN
Query Match      2.9%; Score 31; DB 4; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.00081;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 CATAGCTTGGTGTATGAATTTGGGACTTG 309
|||||
Db 282 CATAGCTTGGTGTATGAATTTGGGACTTG 312
|||||

RESULT 12
BB583163/c
LOCUS BB583163 RIKEN full-length enriched, adult male cecum Mus musculus
DEFINITION cDNA clone 9130204I12 5', mRNA sequence.
ACCESSION BB583163
VERSION BB583163.1 GI:11479707
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 124)
AUTHORS Aizawa,K., Akahira,S., Akimura,T., Arai,A., Arakawa,T.,
Carninci,P., Hanagaki,T., Hayatsu,N., Hiraoka,T., Hirozane,T.,
Hodoyama,Y., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kawai,J.,
Kojima,Y., Konno,H., Kusakabe,M., Matsuyama,T., Miyazaki,A.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y.,
Okido,T., Owa,C., Sakai,C., Sakai,K., Sasaki,D., Sato,K.,
Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
Suzuki,H., Tagawa,A., Takahashi,F., Tanaka,T., Toya,T.,
Watahiki,A., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshiki,A.,
MURAMATSU,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Aizawa,K. et al. 2000)
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S.,

```

Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Thermostabilization and thermoactivation of thermolabile enzymes by  
trehalose and its application for the synthesis of full length  
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
Itoh,M., Kitsuai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,  
Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,  
Okazaki,Y. and Hayashizaki,Y.  
Automated filtration-based high-throughput plasmid preparation  
system. Genome Res. 9 (5), 463-470 (1999)  
Carninci,P. and Hayashizaki,Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
19-44 (1999)  
Please visit our web site (<http://genome.rtc.riken.go.jp>) for  
further details.

#### FEATURES

Location/Qualifiers  
1..124  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="9130204112"  
/sex="male"  
/tissue\_type="cecum"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="RIKEN full-length enriched, adult male cecum"  
/note="Site 1: SalI; Site 2: BamHI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN. Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer [5'  
GAGAGAGAGAGTCCTCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was  
prepared by using trehalose thermo-activated reverse  
transcriptase and subsequently enriched for full-length by  
cap-trapper. cDNA went through one round of normalization  
to Rot = 10.0 and subtraction to Rot = 185.0. Second  
strand cDNA was prepared with the primer adapter of  
sequence [5' GAGAGAGAGATTCAGTGTAATTAATTAATTCCTCCCTCCCTCC  
3']. cDNA was cloned into the XhoI and BamHI sites.  
Vector: a modified pBluescript KS(+) after bulk excision  
from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end:  
BamHI"

#### ORIGIN

Query Match 2.8%; Score 30; DB 2; Length 124;  
Best Local Similarity 100.0%; Pred. No. 0.0027;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1050 TTTCCTTTTGGAAAAAATGAAAAAATGAAAAA 1079  
|||||  
Db 58 TTTCCTTTTGGAAAAAATGAAAAAATGAAAAA 29  
|||||

RESULT 13  
AI591310/c  
LOCUS  
DEFINITION  
tw11f05.x1 NCI CGAP Brn52 Homo sapiens cDNA clone IMAGE:2259393 3'  
similar to contains\_MER22.b3 MER22 repetitive element ;, mRNA  
sequence.  
AI591310 388 bp mRNA linear EST 14-MAY-1999  
AI591310  
AI591310.1 GI:4600358  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 388)  
REFERENCE  
AUTHORS  
NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE  
National Cancer Institute / National Institute of Neurological  
Disorders and Stroke Brain Tumor Genome Anatomy Project  
(CGAP/BRGAP), Tumor Gene Index

#### JOURNAL COMMENT

Unpublished (1998)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael  
R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Insert Length: 1681 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 388  
POLYA=No.

#### FEATURES

Location/Qualifiers  
1..388  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2259393"  
/tissue\_type="tumor, 5 pooled (see description)"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Brn52"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: SalI;  
Site 2: NotI; This library represents the normalized  
version of NCI CGAP Brn35. Cloned unidirectionally.  
Primer: Oligo dT. Average insert size 1.19 kb. Tumor  
types include: meningioma, oligodendroglioma, astrocytoma  
(grade II), medulloblastoma, astrocytoma (grade IV).  
Constructed by Life Technologies."

#### ORIGIN

Query Match 2.8%; Score 30; DB 1; Length 388;  
Best Local Similarity 100.0%; Pred. No. 0.0026;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1050 TTTCCTTTTGGAAAAAATGAAAAAATGAAAAA 1079  
|||||  
Db 89 TTTCCTTTTGGAAAAAATGAAAAAATGAAAAA 60  
|||||

#### RESULT 14 AI718513/c

LOCUS  
DEFINITION  
as56h02.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone  
IMAGE:232755 3' similar to SW:CA11\_RAT P02454 COLLAGEN ALPHA 1(I)  
CHAIN ;, mRNA sequence.

ACCESSION  
AI718513  
AI718513.1 GI:5035769  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 400)  
REFERENCE  
AUTHORS  
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,  
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,  
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,  
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.  
WashU-NCI human EST Project  
Unpublished (1997)

#### TITLE JOURNAL COMMENT

Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Possible reversed clone; similarity on wrong strand  
Seq primer: -40UP from Gibco  
High quality sequence stop: 388.

FEATURES  
source

Location/Qualifiers  
1. .400  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2332755"  
/sex="male"  
/dev stage="adult, age 25"  
/lab\_host="DH10B (phage resistant)"  
/clone\_lib="Barstead colon HPLRB7"  
/notes="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: EcoRI; Site 2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTAGCAATCTCAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [5' AATTCACTAGTAAT 3' and 5' ATTACTAGTG 3'], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed by Bob Barstead."

## ORIGIN

Query Match 2.8%; Score 30; DB 1; Length 400;  
Best Local Similarity 100.0%; Pred. No. 0.0026;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1050 TTTCCTTTTGAAGAAAAA 1079  
|||||  
DB 88 TTTCCTTTTGAAGAAAAA 59

## RESULT 15

BQ383556/c  
LOCUS  
DEFINITION  
BQ383556  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Xenopus laevis (African clawed frog)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
Xenopodinae; Xenopus; Xenopus.  
1 (bases 1 to 417)  
NIH-XCG http://image.llnl.gov/image/html/xenopuslib.info.shtml.  
National Institute of Child Health and Human Development, National  
Cancer Institute, Xenopus Gene Collection  
Unpublished (2002)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
CDNA Library Preparation:  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC)  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
info@image.llnl.gov  
Plate: L1AM1133 row: 0 column: 9  
Seq primer: -21M13 forward primer (ABI).

FEATURES  
source

Location/Qualifiers  
1. .417  
/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="IMAGE:5049056"  
/sex="female"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NICHDXCG.Ovl"  
/notes="Organ: ovary; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 2.0 kb. Constructed by Life  
Technologies."

## ORIGIN

Query Match 2.8%; Score 30; DB 5; Length 417;  
Best Local Similarity 100.0%; Pred. No. 0.0026;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1050 TTTCCTTTTGAAGAAAAA 1079  
|||||  
DB 107 TTTCCTTTTGAAGAAAAA 78

Search completed: September 12, 2005, 17:23:24  
Job time : 3449.77 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: September 12, 2005, 06:32:02 ; Search time 590.713 Seconds  
(without alignments)  
10813.032 Million cell updates/sec

Title: US-10-780-002-1  
Perfect score: 1079  
Sequence: 1 acgcggggaagaagaaaa.....gaaaaaaaaaaaaaaaaa 1079

Scoring table: OLIGO\_NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001s:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003s:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004s:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID       | Description        |
|------------|-------|-------------|--------|-------------|--------------------|
| 1          | 1079  | 100.0       | 1079   | 8 ADA26498  | Ada26498 Tobacco s |
| 2          | 37    | 3.4         | 121    | 12 ADG00027 | Adg00027 Nicotiana |
| C 3        | 27    | 2.5         | 258    | 2 AAQ61367  | Aaq61367 Human bra |
| C 4        | 27    | 2.5         | 321    | 8 ABX54351  | Abx54351 Bovine ES |
| C 5        | 27    | 2.5         | 446    | 6 ABL2609   | AbL2609 Colon ade  |
| 6          | 27    | 2.5         | 552    | 8 ABZ73132  | Abz73132 Rice leaf |
| 7          | 27    | 2.5         | 1795   | 3 AAF16319  | Aaf16319 Human pro |
| 8          | 27    | 2.5         | 1795   | 6 ABL90691  | AbL90691 Human pol |
| 9          | 26    | 2.4         | 342    | 3 AAC79879  | Aac79879 Human sec |
| 10         | 26    | 2.4         | 375    | 3 AAC79873  | Aac79873 Human sec |
| 11         | 26    | 2.4         | 375    | 4 AAF26589  | Aaf26589 DNA encod |
| 12         | 26    | 2.4         | 814    | 5 AAS11362  | Aas11362 Soybean D |
| 13         | 26    | 2.4         | 1099   | 10 ADA13386 | Ada13386 Human int |
| 14         | 26    | 2.4         | 1664   | 2 AAQ73389  | Aaq73389 Rabies vi |
| 15         | 26    | 2.4         | 1870   | 2 AAZ06371  | Aaz06371 Human rec |
| 16         | 26    | 2.4         | 2243   | 3 AAC59812  | Aac59812 Human sec |
| C 17       | 26    | 2.4         | 3820   | 13 ADR85197 | Adr85197 Aspergill |
| C 18       | 26    | 2.4         | 5820   | 13 ADR84610 | Adr84610 Aspergill |
| C 19       | 26    | 2.4         | 20752  | 4 AAK75098  | Aak75098 Human inm |
| 20         | 26    | 2.4         | 98345  | 13 ABD32892 | Abd32892 Human Can |

|      |    |     |     |             |                    |
|------|----|-----|-----|-------------|--------------------|
| C 21 | 25 | 2.3 | 148 | 5 ABV60761  | Abv60761 Human pro |
| C 22 | 25 | 2.3 | 158 | 5 ABV61010  | Abv61010 Human pro |
| C 23 | 25 | 2.3 | 194 | 5 ABV60879  | Abv60879 Human pro |
| C 24 | 25 | 2.3 | 214 | 8 ABX50662  | Abx50662 Bovine ES |
| 25   | 25 | 2.3 | 219 | 6 ABX29070  | Abx29070 cDNA enco |
| 26   | 25 | 2.3 | 219 | 6 ABS68210  | AbS68210 cDNA enco |
| 27   | 25 | 2.3 | 219 | 10 ADC25204 | Adc25204 Human pro |
| C 28 | 25 | 2.3 | 221 | 5 ABV61804  | Abv61804 Human pro |
| 29   | 25 | 2.3 | 244 | 12 ADO41106 | Ado41106 Human cDN |
| 30   | 25 | 2.3 | 245 | 2 AA303925  | Aav30925 Human sec |
| 31   | 25 | 2.3 | 245 | 5 AAF98404  | Aaf98404 3' portio |
| C 32 | 25 | 2.3 | 276 | 12 ADM91205 | Adm91205 Human DNA |
| 33   | 25 | 2.3 | 282 | 11 ADT96541 | Adt96541 Colon can |
| C 34 | 25 | 2.3 | 290 | 5 ABV61419  | Abv61419 Human pro |
| C 35 | 25 | 2.3 | 310 | 12 ADQ18503 | Adq18503 Human sof |
| C 36 | 25 | 2.3 | 341 | 5 ADI71926  | Adi71926 Human ova |
| C 37 | 25 | 2.3 | 341 | 5 ADL37077  | Adl37077 Human ova |
| 38   | 25 | 2.3 | 343 | 4 AAD05605  | Aad05605 Human sec |
| C 39 | 25 | 2.3 | 372 | 6 ABQ58882  | Abq58882 Human col |
| C 40 | 25 | 2.3 | 393 | 5 ABV61180  | Abv61180 Human pro |
| C 41 | 25 | 2.3 | 403 | 4 AAI82336  | Aai82336 Human pol |
| C 42 | 25 | 2.3 | 409 | 9 ACH45796  | Ach45796 Human foe |
| C 43 | 25 | 2.3 | 419 | 8 ABX48773  | Abx48773 Bovine ES |
| 44   | 25 | 2.3 | 425 | 4 AAI92294  | Aai92294 Human pol |
| C 45 | 25 | 2.3 | 435 | 9 ACH20634  | Ach20634 Human adu |

ALIGNMENTS

RESULT 1  
ADA26498  
ID ADA26498 standard; cDNA; 1079 BP.  
XX  
AC ADA26498;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Tobacco salicylic acid binding protein 2 (SABP2) gene.  
XX  
KW salicylic acid-binding protein; SABP2; disease resistance; plant;  
modulator; tobacco; ds; gene.  
XX  
OS Nicotiana tabacum.  
XX  
FH Key Location/Qualifiers  
FT CDS 42..824  
FT /\*tag= a  
FT /product= "salicylic acid binding protein SABP2"  
XX  
PN WO2003016551-A2.  
XX  
XX 27-FEB-2003.  
PD  
PF 16-AUG-2002; 2002WO-US026312.  
XX  
PR 16-AUG-2001; 2001US-0312863P.  
XX  
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.  
XX  
PI Klessig DF, Kumar D;  
DR WPI: 2003-278578/27.  
DR P-PSDB; ADA26499.  
PT New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for  
producing transgenic plants with increased resistance to disease, or for  
screening SABP2 modulators that confer enhanced resistance of plants to  
disease.  
XX  
PS Claim 1; Fig 4; 98pp; English.  
XX  
CC The invention relates to a novel isolated 1079 bp salicylic acid-binding

CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid  
CC protein, sequences that specifically hybridize with the nucleic acid, the  
CC complement of the nucleic acid or a natural allelic variant of the  
CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for  
CC conferring increased resistance to disease in plants, or for screening  
CC modulators of SABP2, which confer increased or enhanced resistance of  
CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also  
CC useful for producing transgenic plants with increased resistance to  
CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful  
CC for studying the molecular mechanisms responsible for the modulation of  
CC salicylic acid-mediated disease resistance in plants. This sequence  
CC represents the gene encoding the SABP2 protein from tobacco plants.  
XX

XX SQ Sequence 1079 BP; 307 A; 204 C; 234 G; 334 T; 0 U; 0 Other;

Query Match 100.0%; Score 1079; DB 8; Length 1079;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1079; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGCGGGGAAAGAAAGAACTAACAGGCATATAAATTCAAATGAAGGAAGAAACACT 60  
DB |||||  
1 ACGCGGGGAAAGAAAGAACTAACAGGCATATAAATTCAAATGAAGGAAGAAACACT 60  
QY 61 TTGTTTTAGTACATGCTGCATGCCATGAGGTTGGAGTTGGTACAGCTTAAGCCACTGC 120  
DB |||||  
61 TTGTTTTAGTACATGCTGCATGCCATGAGGTTGGAGTTGGTACAGCTTAAGCCACTGC 120  
QY 121 TAGAAGCTCAGGCCATAAGGTTACAGCCCTTGATTAGCAGCTCTTGCCACTGATTTGA 180  
DB |||||  
121 TAGAAGCTCAGGCCATAAGGTTACAGCCCTTGATTAGCAGCTCTTGCCACTGATTTGA 180  
QY 181 GAAATATAGAGAGCTTCGACACCTTTATGATTAATTAATTTGCGATGATGGATGG 240  
DB |||||  
181 GAAATATAGAGAGCTTCGACACCTTTATGATTAATTAATTTGCGATGATGGATGG 240  
QY 241 AATCTCTTTACAGAGATGAGAGGTTATATTAGTGGGCATAGTCTTGCTGATGAAT 300  
DB |||||  
241 AATCTCTTTACAGAGATGAGAGGTTATATTAGTGGGCATAGTCTTGCTGATGAAT 300  
QY 301 TGGGACTCTGCTATGGAAGATATCCACAAAGATCTATGCTGCTGTTCTTGGCTGCTT 360  
DB |||||  
301 TGGGACTCTGCTATGGAAGATATCCACAAAGATCTATGCTGCTGTTCTTGGCTGCTT 360  
QY 361 TCATGCTGATTTCTGTTCAACTCTCTCTTTGTTTGGAAACAGATTAATGAGCGGAC 420  
DB |||||  
361 TCATGCTGATTTCTGTTCAACTCTCTCTTTGTTTGGAAACAGATTAATGAGCGGAC 420  
QY 421 CAGCCGAGAAATGGTTGGATCTCAGTTTTTACCATATGTTCCCTGGAAGGCCACTGA 480  
DB |||||  
421 CAGCCGAGAAATGGTTGGATCTCAGTTTTTACCATATGTTCCCTGGAAGGCCACTGA 480  
QY 481 CATCCATGTTTTTGGCCCAAAAGTTCTTGCTCACAAGCTCTACCAAGCTATGCTCTCG 540  
DB |||||  
481 CATCCATGTTTTTGGCCCAAAAGTTCTTGCTCACAAGCTCTACCAAGCTATGCTCTCG 540  
QY 541 AGGATCTTCGATAGCATCATCTTGGTGAGACCAAGCTCTTTGTTTATGGAAGCCTAT 600  
DB |||||  
541 AGGATCTTCGATAGCATCATCTTGGTGAGACCAAGCTCTTTGTTTATGGAAGCCTAT 600  
QY 601 CGAAGCCCAAGTATTTTCAGATGAACGGTTTGGATCAGTGAAGAGAGTTTACATTGCT 660  
DB |||||  
601 CGAAGCCCAAGTATTTTCAGATGAACGGTTTGGATCAGTGAAGAGAGTTTACATTGCT 660  
QY 661 GCATGAGGATAAGGCATACCAAGAAATTCACGCGATGGCAAAATGCAACATTGGTG 720  
DB |||||  
661 GCATGAGGATAAGGCATACCAAGAAATTCACGCGATGGCAAAATGCAACATTGGTG 720  
QY 721 TCACTGAGCAATAGAGATTAAGGTGCTGATCAGATGCAATGCTATCGAGCCCAAA 780  
DB |||||  
721 TCACTGAGCAATAGAGATTAAGGTGCTGATCAGATGCAATGCTATCGAGCCCAAA 780  
QY 781 AACTTTGCGCTCTCTCTTGGAAATGGCCCATAAATACAACTGATCTCTACATTATGCT 840  
DB |||||

DB 781 AACTTTGCGCTCTCTCTTGGAAATGGCCCATAAATACAACTGATCTCTACATTATGCT 840  
QY 841 TCGTCTCAGTCAAGATTTTCAGTGCATGCTGTAATTTTTTTCTATTTTCGACCGGCGC 900  
DB |||||  
841 TCGTCTCAGTCAAGATTTTCAGTGCATGCTGTAATTTTTTTCTATTTTCGACCGGCGC 900  
QY 901 ATAACTGCTTTGGCTATTTTAAGGATTTGCAAGTAAATTTCACTCTTCTAGTGTGAAGGCT 960  
DB |||||  
901 ATAACTGCTTTGGCTATTTTAAGGATTTGCAAGTAAATTTCACTCTTCTAGTGTGAAGGCT 960  
QY 961 TCCCATAAAGGATTTCTGTTTCTCCATTCAGTGTGTTATGTTGAGATACCTTAAAC 1020  
DB |||||  
961 TCCCATAAAGGATTTCTGTTTCTCCATTCAGTGTGTTATGTTGAGATACCTTAAAC 1020  
QY 1021 CGTATCAATTTCTGTAATGAACCTTCTTCTTCTTTTGGAAAAAATTTTTTTTTT 1079  
DB |||||  
1021 CGTATCAATTTCTGTAATGAACCTTCTTCTTCTTTTGGAAAAAATTTTTTTTTT 1079

RESULT 2  
ADG00027

ID ADG00027 standard; cDNA; 121 BP.

AC ADG00027;

XX 26-FEB-2004 (first entry)

XX Nicotiana tabacum variant bright yellow nucleotide sequence SEQ ID:544.  
XX secondary metabolite modulator; biosynthesis; alkaloid; phenylpropanoid;  
KW tobacco; Nicotiana tabacum variant bright yellow; Nicotiana tabacum BY;  
KW plant; gene; ss.

XX Nicotiana tabacum.

XX WO2003097790-A2.

XX 27-NOV-2003.

XX 16-MAY-2003; 2003WO-EP050171.

XX 17-MAY-2002; 2002EP-00076973.

XX 04-JUL-2002; 2002EP-00077674.

XX (VLAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

XX (VTTB-) VTT BIOTECHNOLOGY.

XX Inze DG, Goossens A, Oksman-Caldentey K, Haekkinen ST, Laakso IJ;

XX WPI; 2004-022853/02.

XX New isolated polypeptides and polynucleotides useful for modulating the

XX biosynthesis of secondary metabolites (e.g. alkaloids or

XX phenylpropanoids) in an organism or its cell.

XX Claim 3; SEQ ID NO 544; 140pp; English.

XX The present invention describes an isolated polypeptide that modulates  
CC the production of at least one secondary metabolite in an organism, or  
CC its derived cell. The polypeptide is selected from a polypeptide encoded  
CC by a polynucleotide comprising any of the 871 nucleotide sequences given  
CC in the specification. The polypeptides and polynucleotides are useful in  
CC modulating the biosynthesis of secondary metabolites (e.g. alkaloids or  
CC phenylpropanoids) in an organism or its derived cell. The present  
CC sequence is used in the exemplification of the present invention.

XX Sequence 121 BP; 44 A; 19 C; 31 G; 27 T; 0 U; 0 Other;

Query Match 3.4%; Score 37; DB 12; Length 121;

Best Local Similarity 100.0%; Pred. No. 0.00077;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 628 GGTTCGATCAGTGAAGAGATTTACATTGTGTGCAC 664



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 11, 2005, 22:54:15 ; Search time 4681.27 Seconds  
(without alignments)  
11168.596 Million cell updates/sec

Title: US-10-780-002-1  
Perfect score: 1079  
Sequence: 1 acgcggggaagaaagaaaa.....gaaaaaaagaaaaaa 1079

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hcg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sbs.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES  |       |             |        |                    |
|------------|-------|-------------|--------|--------------------|
| Result No. | Score | Query Match | Length | Description        |
| 1          | 1079  | 100.0       | 1079   | AY485932 Nicotiana |
| 2          | 352.6 | 32.7        | 1077   | AF269158 Citrus si |
| 3          | 332.8 | 30.8        | 961    | AF178576 Rauwolfia |
| 4          | 312.4 | 29.0        | 891    | AF751530 Catharan  |
| 5          | 295.2 | 27.4        | 935    | BT002859 Arabidops |
| 6          | 292.2 | 27.1        | 792    | AX506625 Sequence  |
| 7          | 292.2 | 27.1        | 792    | BT014881 Arabidops |
| 8          | 273   | 25.3        | 870    | AY086590 Arabidops |
| 9          | 272.6 | 25.3        | 932    | AF316127 Arabidops |
| 10         | 271   | 25.1        | 792    | AX506856 Sequence  |
| 11         | 271   | 25.1        | 792    | AY056068 Arabidops |
| 12         | 269   | 24.9        | 792    | CQ806104 Sequence  |
| 13         | 269   | 24.9        | 792    | AX505522 Sequence  |
| 14         | 269   | 24.9        | 792    | BT015031 Arabidops |
| 15         | 269   | 24.9        | 878    | BT014901 Arabidops |
| 16         | 261   | 24.2        | 902    | AK175772 Arabidops |
| 17         | 256   | 23.7        | 792    | BT010429 Arabidops |
| 18         | 251   | 23.3        | 924    | BT004125 Arabidops |
| 19         | 244.8 | 22.7        | 928    | AK117107 Arabidops |

|    |       |      |        |   |           |                    |
|----|-------|------|--------|---|-----------|--------------------|
| 20 | 241.4 | 22.4 | 771    | 8 | BT006227  | Arabidops          |
| 21 | 239.8 | 22.2 | 771    | 6 | AX506583  | Sequence           |
| 22 | 239.4 | 22.2 | 770    | 6 | AX651280  | Sequence           |
| 23 | 234.2 | 21.7 | 928    | 8 | AY046015  | Arabidops          |
| 24 | 229.2 | 21.2 | 808    | 8 | AY142490  | Arabidops          |
| 25 | 228.2 | 21.1 | 777    | 8 | AY093714  | Arabidops          |
| 26 | 225.8 | 20.9 | 944    | 8 | BT010709  | Arabidops          |
| 27 | 223.6 | 20.7 | 911    | 8 | AY058115  | Arabidops          |
| 28 | 223.2 | 20.7 | 783    | 8 | BT012414  | Arabidops          |
| 29 | 215   | 19.9 | 903    | 8 | AY086981  | Arabidops          |
| 30 | 215   | 19.9 | 904    | 8 | AY074858  | Arabidops          |
| 31 | 212.8 | 19.7 | 828    | 8 | AY142031  | Arabidops          |
| 32 | 209.8 | 19.4 | 789    | 8 | AY455313  | Lycopersi          |
| 33 | 208   | 19.3 | 1087   | 8 | BT012867  | BT012867 Lycopersi |
| 34 | 195.2 | 18.1 | 1073   | 8 | AK061058  | Oryza sat          |
| 35 | 186   | 17.2 | 1078   | 8 | HBU40402  | U40402 Hevea bras  |
| 36 | 186   | 17.2 | 1091   | 6 | AS9586    | Sequence 1         |
| 37 | 171.6 | 15.9 | 1021   | 8 | MEAHXNL   | 229091 M.esculent  |
| 38 | 169   | 15.7 | 777    | 6 | CQ829357  | Sequence           |
| 39 | 168   | 15.6 | 795    | 6 | AX654444  | Sequence           |
| 40 | 156.6 | 14.5 | 777    | 6 | CQ829358  | Sequence           |
| 41 | 150.2 | 13.9 | 12360  | 8 | AC003040  | Arabidops          |
| 42 | 148.6 | 13.8 | 92049  | 8 | AC004482  | Arabidops          |
| 43 | 144.6 | 13.4 | 198780 | 8 | ATCHRIV86 | Arabidops          |
| 44 | 144.6 | 13.4 | 206420 | 8 | ATAP21    | Arabidops          |
| 45 | 136   | 12.6 | 3948   | 8 | MEAJ3281  | Manihot e          |

ALIGNMENTS

|           |  |  |      |        |                 |
|-----------|--|--|------|--------|-----------------|
| RESULT 1  | AY485932   | 1079 bp  | mRNA | linear | PLN 07-JAN-2004 |
| LOCUS     | Nicotiana tabacum  | salicylic acid-binding protein 2 mRNA, complete cds. |      |        |                 |
| ACCESSION | AY485932   |  |      |        |                 |
| VERSION   | AY485932.1   | GI:40549302  |      |        |                 |
| KEYWORDS  | Nicotiana tabacum (common tobacco)   |  |      |        |                 |
| SOURCE    | Nicotiana tabacum  |  |      |        |                 |
| ORGANISM  | Nicotiana tabacum  |  |      |        |                 |
| REFERENCE | 1 (bases 1 to 1079)  |  |      |        |                 |
| AUTHORS   | Kumar, D. and Klessig, D.F.  |  |      |        |                 |
| TITLE     | High-affinity salicylic acid-binding protein 2 is required for plant innate immunity and has salicylic acid-stimulated lipase activity   |  |      |        |                 |
| JOURNAL   | Proc. Natl. Acad. Sci. U.S.A. 100 (26), 16101-16106 (2003)   |  |      |        |                 |
| PUBMED    | 14673096   |  |      |        |                 |
| REFERENCE | 2 (bases 1 to 1079)  |  |      |        |                 |
| AUTHORS   | Kumar, D. and Klessig, D.F.  |  |      |        |                 |
| TITLE     | Direct Submission  |  |      |        |                 |
| JOURNAL   | Submitted (24-NOV-2003) BTI, Tower Rd., Ithaca, NY 14853, USA  |  |      |        |                 |
| FEATURES  | Location/Qualifiers  |  |      |        |                 |
| source    | 1..1079  |  |      |        |                 |
|           | /organism="Nicotiana tabacum"  |  |      |        |                 |
|           | /mol_type="mRNA"   |  |      |        |                 |
|           | /cultivar="Xanthi nc [NN]"   |  |      |        |                 |
|           | /db_xref="taxon:4097"  |  |      |        |                 |
|           | 42..824  |  |      |        |                 |
|           | /note="lipase: SAMP2"  |  |      |        |                 |
| CDS       | /codon_start=1   |  |      |        |                 |
|           | /product="salicylic acid-binding protein 2"  |  |      |        |                 |
|           | /protein_id="AA87711.1"  |  |      |        |                 |
|           | /db_xref="GI:40549303"   |  |      |        |                 |
|           | /translation="MKEGKHFLVHGACHGWSYKLPLEAAGHKVTDALDASGT DLKTEELTYDYTLPLMELMESLSDKVKTLVGLSGMNLGAMKVKPKIYAAV FLAAMPDSVHNSFVLQYNERTPAENWLDQFLPYGSPPEPLSPFLAHLK YLCSPEDLALASSLVRPSSLMEDLSKAKYFTDPRFGSKRVIYVCTSDKGIPEEFQ RWDIGNIVGTEAIEIKGADHMAWLCEPQKLCASLLEIAHKYN" |  |      |        |                 |
| ORIGIN    |  |  |      |        |                 |

|                       |     |   |                   |            |              |        |
|-----------------------|-----|---|-------------------|------------|--------------|--------|
| Query Match           |     | 100.0%;   | Score 1079;       | DB 8;      | Length 1079; |        |
| Best Local Similarity |     | 100.0%;   | Pred. No. 5e-260; |            |              |        |
| Matches 1079;         |     | Conservative  | 0;                | Mismatches | 0;           | Indels |
|                       |     |   |                   |            |              | Gaps   |
| Qy                    | 1   | ACGCGGGAGAGAAAGAACTAAACAGGCATATAATTCAAATGAAGAGGAAACACT          | 60                |            |              |        |
|                       |     |   |                   |            |              |        |
| Db                    | 1   | ACGCGGGAGAGAAAGAACTAAACAGGCATATAATTCAAATGAAGAGGAAACACT          | 60                |            |              |        |
|                       |     |   |                   |            |              |        |
| Qy                    | 61  | TTGTTTGTAGTACATGTCATGTCATGCCATGAGAGTTGGAGTTGGTACAGCTAAAGCCACTGC | 120               |            |              |        |
|                       |     |   |                   |            |              |        |
| Db                    | 61  | TTGTTTGTAGTACATGTCATGCCATGAGAGTTGGAGTTGGTACAGCTAAAGCCACTGC      | 120               |            |              |        |
|                       |     |   |                   |            |              |        |
| Qy                    | 121 | TAGAAGTCAGGCCATAGAGTTACAGCCCTTGATTTAGCAGCTCTTGGCAGCTGATTTGA     | 180               |            |              |        |
|                       |     |   |                   |            |              |        |
| Db                    | 121 | TAGAAGTCAGGCCATAGAGTTACAGCCCTTGATTTAGCAGCTCTTGGCAGCTGATTTGA     | 180               |            |              |        |
|                       |     |   |                   |            |              |        |
| Qy                    | 181 | GAATAATAGAGGCTTCGCACATTTATGATTATACCTTTGCCATTTGAGGAGTTGATGG      | 240               |            |              |        |
|                       |     |   |                   |            |              |        |
| Db                    | 181 | GAATAATAGAGGCTTCGCACATTTATGATTATACCTTTGCCATTTGAGGAGTTGATGG      | 240               |            |              |        |
|                       |     |   |                   |            |              |        |
| Qy                    | 241 | AATCTCTTTTACGAGATGAGAGGTTATATTAGTGGGCGATAGTCTTTGGTGGTATGAATT    | 300               |            |              |        |
|                       |     |   |                   |            |              |        |
| Db                    | 241 | AATCTCTTTTACGAGATGAGAGGTTATATTAGTGGGCGATAGTCTTTGGTGGTATGAATT    | 300               |            |              |        |
|                       |     |   |                   |            |              |        |
| Qy                    | 301 | TGGGACTTGTATGGAAGATATCCACAAAGATCATGCTGCTGTTTCTTGGCTGCTT         | 360               |            |              |        |
|                       |     |   |                   |            |              |        |
| Db                    | 301 | TGGGACTTGTATGGAAGATATCCACAAAGATCATGCTGCTGTTTCTTGGCTGCTT         | 360               |            |              |        |
|                       |     |   |                   |            |              |        |
| Qy                    | 361 | TCATGCTGATTTGTTTCAACATCTCTCTTTGTTTGGAAACAGTATATGACGGACGC        | 420               |            |              |        |
|                       |     |   |                   |            |              |        |
| Db                    | 361 | TCATGCTGATTTGTTTCAACATCTCTCTTTGTTTGGAAACAGTATATGACGGACGC        | 420               |            |              |        |
|                       |     |   |                   |            |              |        |
| Qy                    | 421 | CAGCCGAGAAATGTTTGGATACCTCAGTTTTTACCATATATGTTCCCTGAAGACCACTGA    | 480               |            |              |        |
|                       |     |   |                   |            |              |        |
| Db                    | 421 | CAGCCGAGAAATGTTTGGATACCTCAGTTTTTACCATATATGTTCCCTGAAGACCACTGA    | 480               |            |              |        |
|                       |     |   |                   |            |              |        |
| Qy                    | 481 | CATCCATGTTTTTGGCCCAAGTTCTTGGCTCACAAAGCTCTACAGCTATGCTCTCTG       | 540               |            |              |        |
|                       |     |   |                   |            |              |        |
| Db                    | 481 | CATCCATGTTTTTGGCCCAAGTTCTTGGCTCACAAAGCTCTACAGCTATGCTCTCTG       | 540               |            |              |        |
|                       |     |   |                   |            |              |        |
| Qy                    | 541 | AGGATCTTCATTTAGCATCATCATTTGGTGGAGCCAGCTCTTTGTTTATGGAAGCACTAT    | 600               |            |              |        |
|                       |     |   |                   |            |              |        |
| Db                    | 541 | AGGATCTTCATTTAGCATCATCATTTGGTGGAGCCAGCTCTTTGTTTATGGAAGCACTAT    | 600               |            |              |        |
|                       |     |   |                   |            |              |        |
| Qy                    | 601 | CGAAGCCCAAGTATTTACAGATGAACGGTTTGGATCAGTGAAGAGAGTTTACATTTGT      | 660               |            |              |        |
|                       |     |   |                   |            |              |        |
| Db                    | 601 | CGAAGCCCAAGTATTTACAGATGAACGGTTTGGATCAGTGAAGAGAGTTTACATTTGT      | 660               |            |              |        |
|                       |     |   |                   |            |              |        |
| Qy                    | 661 | GCATGAGGATTAAGGCATACAGAGAAATTCAGCGATGGCAATTTGACAACTATGTTG       | 720               |            |              |        |
|                       |     |   |                   |            |              |        |
| Db                    | 661 | GCATGAGGATTAAGGCATACAGAGAAATTCAGCGATGGCAATTTGACAACTATGTTG       | 720               |            |              |        |
|                       |     |   |                   |            |              |        |
| Qy                    | 721 | TCACCTGAAGCAATAGAGATTAAGGTCGTGATCACATGGCAATGCTATCGAGCCCAAA      | 780               |            |              |        |
|                       |     |   |                   |            |              |        |
| Db                    | 721 | TCACCTGAAGCAATAGAGATTAAGGTCGTGATCACATGGCAATGCTATCGAGCCCAAA      | 780               |            |              |        |
|                       |     |   |                   |            |              |        |
| Qy                    | 781 | AACTTTTGGCCCTCTCTTTGGAAATTTGCCATAAATACAACTGATCTTACATTTATGCT     | 840               |            |              |        |
|                       |     |   |                   |            |              |        |
| Db                    | 781 | AACTTTTGGCCCTCTCTTTGGAAATTTGCCATAAATACAACTGATCTTACATTTATGCT     | 840               |            |              |        |
|                       |     |   |                   |            |              |        |
| Qy                    | 841 | TGCTCTCATGTCAAGATTTTACGTGATGCTGTAATTTTTTTCTATTTTTTCACCGCGGC     | 900               |            |              |        |
|                       |     |   |                   |            |              |        |
| Db                    | 841 | TGCTCTCATGTCAAGATTTTACGTGATGCTGTAATTTTTTTCTATTTTTTCACCGCGGC     | 900               |            |              |        |
|                       |     |   |                   |            |              |        |
| Qy                    | 901 | ATAACTGTCTTGGCTTATTTAAGGATTCAGTAAATTTTCACTCTTCTAGTGTGGAAGGCT    | 960               |            |              |        |
|                       |     |   |                   |            |              |        |
| Db                    | 901 | ATAACTGTCTTGGCTTATTTAAGGATTCAGTAAATTTTCACTCTTCTAGTGTGGAAGGCT    | 960               |            |              |        |
|                       |     |   |                   |            |              |        |
| Qy                    | 961 | TCACATAAGGATTTGTTCTCTCTTCTCCATTTCAAGTGTGTTATGTTGAGATACCTTAAAC   | 1020              |            |              |        |
|                       |     |   |                   |            |              |        |
| Db                    | 961 | TCACATAAGGATTTGTTCTCTCTTCTCCATTTCAAGTGTGTTATGTTGAGATACCTTAAAC   | 1020              |            |              |        |
|                       |     |   |                   |            |              |        |

Qy

1021

CGTATCAATTTCTGTAATGAACACTCTCTCTTTCTCTTTTGAAGAAAAA

1079

Db

1021

CGTATCAATTTCTGTAATGAACACTCTCTCTTTCTCTTTTGAAGAAAAA

1079

RESULT 2

AF269158

Citrus sinensis ethylene-induced esterase mRNA, complete cds.

1077 bp

mRNA

linear

PLN 02-AUG-2002

AF269158

AF269158.1

GI:14279436

Citrus sinensis

Citrus sinensis

ORGANISM

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..1077

/organism="Citrus sinensis"

/mol\_type="mRNA"

/db\_xref="taxon:2711"

55..858

/note="EIE"

/codon\_start=1

/product="ethylene-induced esterase"

/protein\_id="AAK58599.1"

/db\_xref="GI:14279437"

/translation="MEEVVGMEERHFLVHGHNHGWKYLKARLVAGGHRVTAVDL

AASGLNKKRIEDVTHFAYSEPLASPAEKVILVHSLGSGVTALALAGDKPCHK

ISVAFVFTAFMPDTHRPSFVLEQSEKMGKDDSLDLOFSCDASNPISHMLPGR

EPLIKIYQLCPPEDELEKMLVRPGSMFIDNLSEKSFDEGYSKVRVLYCEEDI

GLPKQFQHWMIQNTYFVNEVMEIKGDHWMULSDFOKCDCLSLKTA"

CDS

ORIGIN

Query Match

32.7%;

Score 352.6;

DB 8;

Length 1077;

Best Local Similarity

65.8%;

Pred. No. 1.2e-77;

Matches 530;

Conservative

0;

Mismatches

269;

Indels

6;

Gaps

1;

Qy

19

AACTAACAGGCGATAAAATTCAAATGAAGAGGAAAGAAACACTTTCTTTTAGTACATGGT

78

Db

47

AAAAGAAAATGGAAGAGTAGTAGCATGGAAGAGAGCAATTTCTTAGTTCTATGGAG

106

Qy

79

CATGCCATGAGGTTGGAGTTGGTACAGCTAAAGCCACTGCTAGAGCTGCAGGCCATA

138

Db

107

TAAACCATGAGGAGCATGGTTGGTACAACTGAAGAGCAAGGCTGGTGGGGGGTCA

166

Qy

139

AGGTTTACAGCCCTTCATTTTAGCAGCTTCTGGCAGCTGATTTGAGAAAAATAGAGGAGCTTC

198

Db

167

GGGTGACGGCTGTGACCTAGCCGCTCGGCGATCAACATGAGAGAAATGAGGATGTGC

226

Qy

199

GCACATTTATGATTTATCTTTGCCATTTAGTGGAGTTGATGGAATCTCTTTTCAGCAGATG

258

Db

227

ACACATTCATGCAACAGTGAAGCTTGTGAGGTTTGGCATCACTTCTCTGCGGAAG

286

Qy

259

AGAGGTTATATTAGTGGGCATAGTCTTGGTGGTATGAAATTTGGAATTGATGGA

318

Db

287

AAAAGGTCATCTTCTCGGACACAGTCTTGGGGGGCGTCACTTGGCCCTTGC

346

Qy

319

AGTATCCACAAAAGATCTATGCTGCTGTTTCTTGGCTGCTTTCATGCTGATTTCTGTT

378

Db 347 AATTCACACAAAAATCTCCGTGGCTGTTTTCGTAACTGCATTCATCCCTGCACACACAC 406

QY 379 ACAACTCTCTCTTTGTTGGAACAGTATTAATCAGCGGAGCCCA-----GCCGGAATTT 432

Db 407 ACCGGCCATCTTTGTTTGGACAGATTTCTGAGAAGATGGGAAAAGAGGACGACAGCT 466

QY 433 GGTGGATACTAGTTTTCACCATATGTTTCCCTCGAAGAGCCACTGACATCATCATGTTT 492

Db 467 GGTGGACACTCAATTTTCAATGTGACGCGTCAAAATCCATCTCACATTTCCATCTCTTT 526

QY 493 TTGGCCCAAGTTCTTGGCTCACAGCTCTACAGCTATGCTCTCTGAGGATCTTGGCAT 552

Db 527 TCGGCCCGCAGTTCTTTGACTATCAAGATCTATCAGCTTTGTCTCTCTGAGGATCTGGAGC 586

QY 553 TAGCATCATCATTTGGTGAGACCAAGCTCTTTGTTTATGGAAGACCTATCGAAGGCCAAGT 612

Db 587 TGCCCAAGATGTTGGTGAGCCAGGATCAATGTTTATAGCAACTTATCGAAGGAAGTA 646

QY 613 ATTTCAAGATGAACGGTTTGGATCAGTGAAGAGAGTTTACATTTGTGTCACCTGAGGATA 672

Db 647 AGTTACGCGATGAAGATACGGATCTGTTAAGCGAGTTTATCTTTGTATGCGAAGAGGATA 706

QY 673 AAGCATACCAAGAAATTCACGAGTGGCAATTCACAACATTTGTGTCTACCTGAAGCAA 732

Db 707 TTGGTCTCCCTAAGCAATTTACGACTGGATGATCCAAAACCTATCCCGTTAATGAAGTGA 766

QY 733 TAGAGATTAAGGTGCTGATCAGATGCAATGCTATGCGAGCCGCCAAAACCTTTGCGCCT 792

Db 767 TGGAGATCAAGGCGGTGATCAATGCGAATGCTTTCCGACCCACAGAAACCTTTGCGATT 826

QY 793 CTCTCTTGGAAATGGCCCATMAATA 817

Db 827 GTCTGCTCAGATTTCTCTTAAGTA 851

RESULT 3

AF178576

LOCUS

DEFINITION

AF178576

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

AF178576 961 bp mRNA linear PLN 29-MAR-2000

Rauvolfia serpentina polynuridine aldehyde esterase mRNA, complete cds.

AF178576

AF178576.1 GI:6651392

Rauvolfia serpentina (serpentwood)

Rauvolfia serpentina

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Gentianales; Apocynaceae; Rauvolfioideae; Vinceae; Rauvolfia.

1 (bases 1 to 961)

Dogru, E., Warzecha, H., Seibel, F., Haebel, S., Lottspeich, F. and Stockigt, J.

The gene encoding polynuridine aldehyde esterase of monoterpene indole alkaloid biosynthesis in plants is an ortholog of the alpha/betahydrolase super family

Eur. J. Biochem. 267 (5), 1397-1406 (2000)

20156243

10691977

2 (bases 1 to 961)

Warzecha, H., Dogru, E. and Stockigt, J.

Direct Submission

Submitted (18-AUG-1999) Institut fur Pharmazie, Johannes Gutenberg Universitaet, Staudinger Weg 5, Mainz 55099, Germany

Location/Qualifiers

1. .961

/organism="Rauvolfia serpentina"

/mol\_type="mRNA"

/db\_xref="taxon:4060"

/note="cell suspension culture"

15. .809

/function="catalyzes 16-epi-velllosimine formation from polynuridine aldehyde"

/function="involved in ajmaline biosynthesis"

/codon\_start=1

/product="polynuridine aldehyde esterase"

/protein\_id="AAF22288.1"

/db\_xref="GI:6651392"

/translation="MHSAAAKQKHFLVHGGCLGANIWKLPKLLSAGAHKTVADVLSAAGINRRRLDEIHTFDYSEPLMEVNASIPPEDEKVLHLSFGMSLGLAMETYPEKISAVAFMSAMPDPNHSILTYPEKYNEKCPADMLDSQFSTYGPENFGMSMILGPQFMALKMFONCSVEDLELAKMLTRPGSLFQDLAKAKKFSERYGSKRAYIFCNEDKSFPVEFKWQFVESVGADKVKEIKEADHMGMLSQPREVCKLLDISDS"

ORIGIN

Query Match 30.8%; Score 332.8; DB 8; Length 961;

Best Local Similarity 64.3%; Pred. No. 1.1e-72;

Matches 499; Conservative 0; Mismatches 277; Indels 0; Gaps 0;

QY 45 AAGGAGGAAACACATTTGTTTGTACATGTCGATGTCATGAGGTTTGGAGTTGGTAC 104

Db 36 AAGCAACAAACATTTTGTCTGTACACGCGGATGTCTCGAGCTTGGATCTGGTAC 95

QY 105 AAGCTAAAGCCACTGCTAGAACCTGCAGGCCATAAGGTTTACAGCCCTTGTATTTAGCAGCT 164

Db 96 AAGCTCAAGCCGCTGCTCGAGTCAGCCGACATAAGGTACCCGCTTGACCTGTCGGCC 155

QY 165 TCTGCACTGATTTGAGAAAAATAGAGAGCTTGCACACATTTATGATTTATATTGTTGCCA 224

Db 156 GCGCGCATCAACCCAAAGAGGCTCGATGAGATTTCACATTTTCGGGACTACTCGGAGCCC 215

QY 225 TTGATGGAGTTGATGGAATCTCTTTTCAGCAGATGAGAGGTTTATATTAGTGGGCGCAT 284

Db 216 TTGATGGAGTCTGCTGATGTTTCTCTGATGAGAGGTTTGTCTCTTCTGGCCATAGC 275

QY 285 CTGTTGTTGATGAATTTTGGGACTTCTGATGAAAAAGTATCCACAAAAGATCTATGCTGCT 344

Db 276 TTTGGTGGCATGAGTTTGGCTCTTGCCATGGAACTCTCCAGAGAGATATCAGTTGCT 335

QY 345 GTTTTCTTTGGCTGCTTTCATGCTGATTTCTGTTTCACAACTCTCTCTCTTTTGGAAACAG 404

Db 336 GTTTTATGTTGCAATGATGCTCTGATCTTAACCACTCACTAACTTATCCGTTTGAGAAG 395

QY 405 TATATAGCGGACGCCAGCCAGAAATTTGGTTGGATCTCAGTTTGTATGAGGCTCAAAATGTTTC 464

Db 396 TACAATGAGAGTGTCCGCGAGATATGATGTTGGACTCACAGTTTTCACCTACGGAAC 455

QY 465 CTGGAAGAGCCACTGCACATCCATGCTTTTGGCCCAAGTCTTGCTGCACAAAGCTCTAC 524

Db 456 CCAGAGACCCAGGAATGTCAATGATTTGGAAGTCTGATGAGGCTTATGAGGCTCAAAATGTTTC 515

QY 525 CAGCTATGCTCTCTGAGGATCTTGCATTTAGCATCATCATTTGTTGAGACCAAGCTCTTTG 584

Db 516 CAGAATGCTCAGTCAGGACCTTGAATTAGCCAAAATGTTGACTCGACCCAGGTTCTGTTA 575

QY 585 TTTATGGAAGACCTATCGAAGCCCAAGTATTTTCAGATGAACGGTTTGGATCAGTGAAG 644

Db 576 TTTTTCGAAGATTTGGCCAAAGCCAAAAGTTTCTCAACCGAGAGGTACGGTTTCGGTGAAG 635

QY 645 AGAGTTTACATTTGTCGCACTGAGCATAAAGCCATACCAGAGAAATCCAGCGATGGCAA 704

Db 636 CGAGCTTATATCTTTTGCAATGAAGTAAATCATTTCCAGTTGAGTTTCAGAAATGTTT 695

QY 705 ATTGACAACATTTGGTGTCTCACTGAACAATAGAGATTTAAAGGTGTGATCAGATGGCAATG 764

Db 696 GTTGAAGTGTGGAGCTGATAAAGTAAAGAAATCAAGAAGCAGATCATATGGGAATG 755

QY 765 CTATCGGAGCCCAAAAACCTTTGCGCCTCTCTCTTTGGAAATTTGCCCAATAAATACAA 820

Db 756 CTTTCGACGCAAGGGAAGTTTGCAGGTGCTGCTTGTATATATCAGATTTCATAAAA 811

RESULT 4

AY751530

LOCUS

DEFINITION

ACCESSION

AY751530

Catharanthus roseus protein S mRNA, complete cds.

AY751530

|            |  |  |
|------------|--|--|
| VERSION    | AY751530.1   | GI:53830669  |
| KEYWORDS   | Catharanthus roseus (Madagascar periwinkle)  |  |
| SOURCE     | Catharanthus roseus  |  |
| ORGANISM   | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Gentianales; Apocynaceae; Rauvolfioideae; Vinceae; Catharanthus.   |  |
| REFERENCE  | 1 (bases 1 to 891)   |  |
| AUTHORS    | Lemenager,D., Rideau,M. and Clastre,M.   |  |
| TITLE      | Cloning of the Catharanthus roseus protein S (CrPS) associated with monoterpenoid indole alkaloid production   |  |
| JOURNAL    | Unpublished  |  |
| REFERENCE  | 2 (bases 1 to 891)   |  |
| AUTHORS    | Lemenager,D., Rideau,M. and Clastre,M.   |  |
| TITLE      | Direct Submission  |  |
| JOURNAL    | Submitted (15-SEP-2004) Laboratoire de Biologie Molculaire et Biochimie Vegetale, Faculte de Pharmacie, 31 Avenue Monge, Tours 37200, France   |  |
| FEATURES   | Location/Qualifiers  |  |
| source     | 1..891   |  |
|            | /organism="Catharanthus roseus"  |  |
|            | /mol_type="mRNA"   |  |
|            | /db_xref="taxon:4058"  |  |
|            | 48..824  |  |
|            | /note="hydrolase; CrPS, alpha/beta fold family protein"  |  |
|            | /codon_start=1   |  |
|            | /product="protein S"   |  |
|            | /protein_id="AAU95203.1"   |  |
|            | /db_xref="GI:53830670"   |  |
|            | /translation="MEVVKHFVTHGVGHGAWVYKLPRTAAGHRCITAVNLAAGSI NEKKUEVRSSIDYAPLLEVDVSPENEKVLVGHSGGQMTAAVGMKFNKSLAV FLNAMPDTEPNSVLYEYATKTPPEAKWQCFQSAYGDPPTISLVCGPEFISSTLYH LSPIDHALGKTLVRPGLFIEDLLKAEKFTBEGFGSPRVVYVIAEDKTIPTPEFQRW MTENNPVKEVKEIKGADHMPFSPKDELSQLLDIAKXHA"   |  |
| ORIGIN     | Query Match 29.0%; Score 312.4; DB 8; Length 891;<br>Best Local Similarity 63.4%; Pred. No. 1.5e-67;<br>Matches 497; Conservative 0; Mismatches 281; Indels 6; Gaps 1;   |  |
| QY         | 31   | ATAAAATCCAATGAAGGAGGAAACACTTTGTTTGTACATCTGTGATGCCATGGAG 90         |
| DB         | 37   | AAAAAANAATGGAGTAATGAACCATTTTGTACGCTCCACGGCTGAGCCAGGAG 96           |
| QY         | 91   | GTTGGAGTTGGTACAGCTAAAGCCACTGCTAGAAGCTGCAGGCCCAATGAAGTTCAGAGCCC 150 |
| DB         | 97   | CTTGGGTCTATTACAGCTGAAACCTAGAAATTGAGGCAGCAGCCACCGTTGTACGGCGG 156    |
| QY         | 151  | TTGATTAGCAGTTCTGGCACTGATTTGAGAAAATAGAGAGCTTGCACACTTATG 210         |
| DB         | 157  | TTAAATTTAGCTGTTCTGGAATCAACGAAAGAAACTCGAAGAAGTTTCGGTCTTCAATTG 216   |
| QY         | 211  | ATTATACTTTGCCATTGAGGTGATGCAATCTCTTTTTCACAGATGAGAGGTTATAT 270       |
| DB         | 217  | ATTACCGCCGCGTGTGTGGAGTTTGGATTTGTTCTGTAATGAGAGGTAATAC 276           |
| QY         | 271  | TAGTGGGGCATAGTCTTGGTGGTATGAATTTGGGACTTGTCTATGGAAAAGTATCCACAAA 330  |
| DB         | 277  | TGTTAGACATAGCGCGCGTGGTATGACGCGCGCTGTTGGGATGGAGAAATTTCCGAATA 336    |
| QY         | 331  | AGATCTATGCTGTTGTTCTTGGCTGCTTTCATGCTGATTTCTGTTTCACACTCTCTCT 390     |
| DB         | 337  | AGATCTCTTTAGCTGCTCTTTTGAATGCTATAATGCTGTATACCGGAAAACAGACCATCTT 396  |
| QY         | 391  | TTGTTTTGGAACAGTATAATGAGCGGACCGCAGCGGAGAAATGGTTGGATCTCATGTTT 450    |
| DB         | 397  | ATGTTTTGGAAGGTACACTGCAAGACTCCACAGAACATGGAAGGATTCGCGATTCT 456       |
| QY         | 451  | TACCATATGTTTCCCTGAAGAGCCACTGCATCCCATGTTTGTGGCCCAAGTCTTGG 510       |
| DB         | 457  | CAGCCTATGGAGATCCT-----CCTATTACATCTTTGGTCTGCGGCTCTGAGTTCAATT 510    |
| QY         | 511  | CTCAAGCTCTACAGCTATGCTCTCTGAGGATCTTGCAATTAGCATCATCATTTGGTGA 570     |
| DB         | 511  | CATCCTCTGTATCACCCTCTCCCTTATTGAGGATCATGCAATGGGAAAAATATTGGTAA 570    |
| QY         | 571  | GACCAAGCTCTTTGTTTATGGAAGACCTTATCGAAGGCCAAGTATTTTACAGATGAACGGT 630  |
| DB         | 571  | GGCCAGGCTCACTATTTATAGAGGATTTATTGGAAGGCAGAAAAAATTTACAGAGGAAGAT 630  |
| QY         | 631  | TTGGATCAGTGAAGAGAGTTTACATTTGTGTCCACTGAGGATAAAGGCATACCAAGAGAT 690   |
| DB         | 631  | TTGGATCAGTACCAAGAGTTTATGTAATAGCTGTCTGAGATAAAACATATACCACCTGAT 690   |
| QY         | 691  | TCCAGCGATGCAAAATGACAACTTGGTGTCACTGAAGCAATAGAGATTTAAAGTGTCTG 750    |
| DB         | 691  | TCCAAGTGTGATGATCGAAAACCAACCCAGTTAGGAAGTCAAGAGATCAAGGTGTCTG 750     |
| QY         | 751  | ATCATATGCCATGCTATGCGAGGCCCAAAACCTTTGCGCCTCTCTCTTTGGAAAATTGCCC 810  |
| DB         | 751  | ATCATATGCCATGCTTCTTAACCTGATGAATATCCCAATGCTTTTGGCATATAGCCA 810      |
| QY         | 811  | ATAA 814   |
| DB         | 811  | AGAA 814   |
| RESULT 5   | BT002859 935 bp mRNA linear PLN 15-JAN-2003  |  |
| LOCUS      | Arabidopsis thaliana clone RAFL14-80-F10 (R20356) putative   |  |
| DEFINITION | acetone-cyanohydrin lyase (At2g23620) mRNA, complete cds.  |  |
| ACCESSION  | BT002859   |  |
| VERSION    | BT002859.1   |  |
| KEYWORDS   | FLI_CDNA.  |  |
| SOURCE     | Arabidopsis thaliana (thale cress)   |  |
| ORGANISM   | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.   |  |
| REFERENCE  | 1 (bases 1 to 935)   |  |
| AUTHORS    | Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Heuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A. |  |
| TITLE      | Arabidopsis Full Length cDNA Clones  |  |
| REFERENCE  | 2 (bases 1 to 935)   |  |
| AUTHORS    | Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Heuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A. |  |
| TITLE      | Direct Submission  |  |
| JOURNAL    | Submitted (15-JAN-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA   |  |
| COMMENT    | RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA'); Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.   |  |
|            | The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Heuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.       |  |

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.

#### FEATURES

source  
1. .935  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="RAF14-80-F10 (R20356)"  
/ecotype="Columbia"  
/note="This clone is in a modified pBluescript vector2 (lambda PS) as a BamHI/XhoI insert."  
gene  
1. .935  
/gene="At2g23620"  
5'UTR  
1. .43  
/gene="At2g23620"  
44. .835  
/gene="At2g23620"  
/codon\_start=1  
/evidence=experimental  
/product="putative acetone-cyanohydrin lyase"  
/protein\_id="AAO22676.1"  
/db\_xref="GI:27754457"  
/translation="MSEERKQHFVLVHSGCHGAWCYKVLLEAVGHRVAVDLAA SGIDITRSDIPTQSEPLTKLLTSLPNDKLVGLVHSGHGLNLAIAEMKFPKKI SVAVFLTAPMDEHSPFVLDKFGSNMPEAWMGTEPEFYDSDNSGLSMFSPDFMK LGLYQLSPLELDLELLMRPGSLFINLDSKMKNSDEGYSVPRFIVCKEDKAIFE ERQRWIDNFPVNLVMEETDHPMFCKPQQLSDYFLKIADKFV"  
misc\_difference  
341  
/gene="At2g23620"  
/note="compared to genomic sequence resulting in an amino acid sequence difference"  
/replaces="g"  
836. .935  
/gene="At2g23620"

#### ORIGIN

Query Match 27.4%; Score 295.2; DB 8; Length 935;  
Best Local Similarity 62.7%; Pred. No. 3.1e-63;  
Matches 494; Conservative 0; Mismatches 288; Indels 6; Gaps 2;  
QY 53 AAAACACTTTTGTAGTACATGGTCATGCCATGCCAGGTGGAGTTGGATGGTACAGCTAAA 112  
DB 64 ACAACATTTTGTACTAGTACATGGTTCTGTCATGCCATGCCAGGTGGTGTGTGTACAAAGTTAA 123  
QY 113 GCCACTGTAGAGCTGCAGGCCATAAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCAC 172  
DB 124 GCCGCTGTAGAGCGCGTGGCCACCGCGTAATCTGTGTGACCTTACCTCCCGAAT 183  
QY 173 TGA---TTTGAGAAAATAGAGGAGCTTCGCACACTTTATGATATATCTTGGCAATTGAT 229  
DB 184 AGACACAACAGGTCGATCACTGACATCCCATCCACATCGCAACAATCTCGAGCCATTGAC 243  
QY 230 GGAGTTGATGGAATCTCTTTCAGCAGATGAGAAGTTATATAGTGGGCGCATAGCTTGG 289  
DB 244 GAAGCTCCTGACCTCATTTGCCAAATGATGAAAAGGTTGTGCTGTTGGTGCACAGCTTTGG 303  
QY 290 TGGTATGAATTTGGGACTTGTCTATGAAAAGTATCCACAAAAGATCTATGCTGCTGTTT 349  
DB 304 TGGCTTGAACCTTAGCCATAGCCATGGAAGAAGTTTCCAAAANAATCTCTGTGCTGATT 363  
QY 350 CTTGGCTGCTTTTCATGCTGATTCTGTTCACAACTCCTCCTTTGTTTGGAAACAGATATA 409  
DB 364 CTTGACTGCTTTTCATGCCGACACCGCAACTCACCACCTCTTCGCTTCTGCAAGTTGG 423  
QY 410 TGAGCGGACGCCAGCCAGAAATTGGTTGGATACTCAGTTTTCACCATATGGTTCCCTGCA 469  
DB 424 AAGCAACATGCTCAAGAAGCATGGATGGCGACCGGAATTCGAACCTTATGGTTACAGCAA 483

QY 470 AGAGCCACTGACATCCATGTTTTTTGGCCCAAAGTTCTTGGCTCACAAGCTCTACAGCT 529  
DB 484 TTCCGGACTGA---GTATGTTTTTTAGCCCTGACTTTCATGAAGTTGGGTCTCTACAGCT 540  
QY 530 ATGCTCTCTGAGGATCTTGCATTTAGCATCATTTGGTGAGACCAAGCTCTTTGTTTAT 589  
DB 541 TTCTCCAGTTGAGGATCTTGAACCTGGGATTTACTTTTAATGAGGCCAGGATCGTTATTTAT 600  
QY 590 GGAAGACCTATCGAAGGCCAAGTATTTTACAGATCAACCGTTTGGATCAGTGAAGAGCT 649  
DB 601 TTAACGATTTATCGAAGATGAANAATCTTCTCGGATGAGGATATGGTCTGTTCTCTCGAT 660  
QY 650 TTACATTTGTGCTGACTGAGGATAAAGGCATACAGAGAATTTCCAGCGATGGCAATGGA 709  
DB 661 TTTTCATAGTGTGTAAGAGGACAAAGCAATTCACAGAGAACGCCAGAGATGATGATGCA 720  
QY 710 CAACATTTGGTGTCACTGAGCAATAGAGATTTAAAGGTGCTGATCATGCGCAATGCTATG 769  
DB 721 TAAATTTTCGGTGAATTTAGTGATGGAGATGGAGGAGACAGATCATATGCGCAATGTTCTG 780  
QY 770 CGAGCCCCCAAAACTTTTGGCCTCTCTTGGAAATTCGCCATAAATACAACTGATCTCT 829  
DB 781 CAAGCCTCAGCAACTCAGTGATTACTTCTGAAAATCGCGACAATTCGTTTAATCAAA 840  
QY 830 ACATATG 837  
DB 841 TCTTCATG 848  
RESULT 6  
AX506625  
LOCUS AX506625  
DEFINITION Sequence 1320 from Patent WO216655.  
ACCESSION AX506625  
VERSION AX506625.1 GI:23387862  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
REFERENCE  
1  
AUTHORS Harper, J.F., Krebs, J., Wang, X. and Zhu, T.  
TITLE Stress-regulated genes of plants, transgenic plants containing same, and methods of use  
JOURNAL Patent: WO 0216655-A 1320 28-FEB-2002;  
The Scripps Research Institute (US); Syngenta Participations AG (CH)  
FEATURES  
source  
1. .792  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"  
ORIGIN  
Query Match 27.1%; Score 292.2; DB 6; Length 792;  
Best Local Similarity 63.1%; Pred. No. 1.8e-62;  
Matches 485; Conservative 0; Mismatches 278; Indels 6; Gaps 2;  
QY 53 AAAACACTTTTGTAGTACATGGTGCATGCCATGCCAGGTGGAGTTGGTACAAAGCTAAA 112  
DB 21 ACAACATTTTGTACTAGTACATGGTTCTGTCATGCCATGCCAGTGTGCTGTGTACAAAGTTAA 80  
QY 113 GCCACTGTAGAGCTGCAGGCCATAAGTTACAGCCCTTGATTTAGCAGCTTCTGGCAC 172  
DB 81 GCCGCTGTAGAGCGCGTGGGCCACCGCGTAATCTGTGGAATTTAGCTCCCGAAT 140  
QY 173 TGA---TTTGAGAAAATAGAGGAGCTTCGCACACTTTTATGATATATCTTTGCCATTGAT 229  
DB 141 AGACACAACAGGTCGATCACTGACATCCCAKATCGCAACATCTCGAGCCATTGAC 200  
QY 230 GGAGTTGATGGAATCTCTTTCAGCAGATGAGAAGTTATATTTAGTGGGCGCATAGTCTTGG 289

Db 201 GAAGCTCCTGACCTCATTGCCAAATGATGAAAGGTTGTGCTCGTTGGTCACAGCTTTGG 260  
QY 290 TGGTATGAATTTGGGACTTGTATGGAAGATATCCAAAAGATCTATGCTGCTGTTTT 349  
Db 261 TGGCTTGAATTTAGCCTAGCCTAGCAAGTTTCCGAAAAAATCTCTGTGCTGTATT 320  
QY 350 CTTGGCTGCTTTCATGCTGATCTGTTCCAACTCTCTCTTTGTTGGACATATAA 409  
Db 321 CTTGACTGCTTTTCATGCCCGACACCGAACCTCACCATCTCTGCTGTGGACAAGTTGG 380  
QY 410 TGAGCGGACCGCAGCGGAAATTTGGTTGGATATCTAGTTTACCATATGTTTCCCTCA 469  
Db 381 AAGCAACATGCTCAGAGCATGATGGCAGCGAATTCGAACCTTATGTTTCAGACAA 440  
QY 470 AGAGCCACTGACATCCATGTTTTTGGCCCAAGTTCTTTGGTCCAAAGCTCTACAGCT 529  
Db 441 TTCCGGACTGA--GTATGTTTTTATGCTGCTGACTGATGAAGTTGGGTCTCTACAGCT 497  
QY 530 ATGCTCTCTGAGGATCTTGCATTAGCATCATCTTGGTGAGACCAAGCTCTTTGTTAT 589  
Db 498 TTCTCCAGTTGAGGATCTTGAATCTGGATTTCTTTAATGAGGCCAGGATCGTTATTAT 557  
QY 590 GGAAGACCTATGAAAGCCCAAGTATTTACAGATGAACGGTTTGGATCAGTGAAGAGAT 649  
Db 558 TAACGATTTATCGAGATGAAATCTTCTCGGATGAAGATATGGTCTGTTCTCGAGT 617  
QY 650 TTACATTTGTGACCTGAGGATAAGGCATATCAGAAAGTATCCAGCGATGGCAATTGA 709  
Db 618 TTTCTAGTGTGTAAGAGGACAAAGCAATTCAGAAAGAACGCCAGAGATGGATGTA 677  
QY 710 CAACATTTGTGCTGACGTAGCAATAGAGATTAAGGTGCTGATCATCGCAATGCTATG 769  
Db 678 TAAATTTCCGGTGAATTTAGTATGAGATGAGGAGACAGATCATATGCCAATGTTCTG 737  
QY 770 CGAGCCCCAAAACCTTTGGCCCTCTCTTTGAAAATTGCCATAAATAC 818  
Db 738 CAAGCCTCAGCACTCAGTGATTACTTCTGAAATTCGGGACAAATTC 786

RESULT 7  
LOCUS BT014881 792 bp mRNA linear PLN 05-JUN-2004  
DEFINITION Arabidopsis thaliana At2g23620 gene, complete cds.

ACCESSION BT014881  
VERSION BT014881.1 GI:48310670  
KEYWORDS FLI CDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 792)  
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Shinn,P. and Ecker,J.R.  
TITLE Arabidopsis ORF clones  
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 792)  
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Shinn,P. and Ecker,J.R.  
TITLE Direct Submission  
JOURNAL Submitted (05-JUN-2004) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

FEATURES  
source Location/Qualifiers  
1..792  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="U67545"  
/ecotype="Columbia"  
/notes="This clone is in pUNI 51"  
1..792  
/note="unknown protein"

CDS

/codon\_start=1  
/product="At2g23620"  
/protein\_id="AAI41864.1"  
/db\_xref="GI:48310671"  
translation="MSEKRRQHFVLVHSGHCAGWCYKVKPLLEAVGHRVTAVDLAA  
SGIDTTSITDIPTEQVSEPLTLKLSLPNDEKVLVLVHSGHFGGLNLAIAMEKPEKI  
SVAVPLTAFMPDPTBHSFVLDFKFSNNPDEAWNGTEPEPYGSDNSGLSMFSPDFMK  
LGLYLQSPVEDLELGLLMPGSLPINDLSKMKNFSDEGYSVPRVFIUCKEDRAIPE  
ERQRWMDINFFVNLVMEETDHPMFCPKPQOLSDFLKLADKFV"

ORIGIN

Query Match 27.1%; Score 292.2; DB 8; Length 792;  
Best Local Similarity 63.1%; Pred. No. 1.8e-62;  
Matches 485; Conservative 0; Mismatches 278; Indels 6; Gaps 2;  
QY 53 AAAACACTTTGTTTGTAGTACATGTGTCATGCGAGGTTGGAGTTGGTACAAGCTAAA 112  
Db 21 ACAACATTTTGTACTAGTACATGTTCTGTCATGCGCGCTGGTCTGTTCAAGGTTAA 80  
QY 113 GCCACTGTAGAACTGCAGGCCATAGGTTTACAGCCCTTGATTTAGCAGCTTCTGGCAC 172  
Db 81 GCCGCTGTAGAGCGGTGGGCCACCGCGTAACCTGCTGTGGACTTAGTCTCCGGAAT 140  
QY 173 TGA---TTTGAGAAAAATAGAGGAGCTTTCGCACACTTTATGATTATATCTTTGCCATTGAT 229  
Db 141 AGACACAAGAGGTGCTGATGACATCCCAATCGCAACAATCTCGAGGCCATTGAC 200  
QY 230 GGAGTTGATGAATCTCTTTTCAGCAGATGAGAGGTTATATAGTGGGGCATAGTCTGG 289  
Db 201 GAAGCTCCTGACCTCATTTGCCAAATGATGAAAAGGTTGTGCTCGTTGTCACAGCTTTGG 260  
QY 290 TGGTATGAATTTGGGACTTGTCTATGGAAGATATCCACAAAGATCTATGCTGCTGTTT 349  
Db 261 TGGCTGAACCTTAGCCATAGCCATGGAAGAGTTTCCCGAAAAAATCTCTGCTGCTGTAAT 320  
QY 350 CTTGCTGCTTTTCATGCTGATTTCTGTTCACAACTCTCTCTCTTTGTTGGAAACAGTATAA 409  
Db 321 CTTGACTGCTTTTCATGCGGACACCGAATCTCACCATCTTCTGCTTTGGACAGTTGG 380  
QY 410 TGAGCGGACCGCAGCGAGAAATGTTGGATCTCATTTGTTGAGACCAAGCTCTTTGTTAT 469  
Db 381 AAGCAACATGCTCAGAAAGCATGATGGGCAACCGAATTCGAACCTTATGTTTCAGACAA 440  
QY 470 AGAGCCACTGACATCCATGTTTGTGGCCCAAGATTTCTTGGCTCACAGCTCTACAGCT 529  
Db 441 TTCCGGACTGA--GTATGTTTTTATGCTGCTGACTTTCATGAAGTTGGGTCTCTACAGCT 497  
QY 530 ATGCTCTCTCGAGGATCTTGTGATTAGCATCATTTGTTGAGACCAAGCTCTTTGTTAT 589  
Db 498 TTCTCCAGTTGAGGATCTTGAATCTGGAATTTCTTTAATGAGGCCAGGATCGTTATTAT 557  
QY 590 GGAAGACCTATGAAAGCCCAAGTATTTCAAGATGAACGGTTTGGATCAGTGAAGAGAT 649  
Db 558 TAAACGATTTATCGAAGATGAAAACTTCTCGGATGAAGGATATGGGTCTGTTCTCGAGT 617  
QY 650 TTACATTTGTGCACTGAGGATAAGGCATACCAAGAAATTCACAGCATGCAAAATGA 709  
Db 618 TTTCTAGTGTGTAAAGAGGACAAAGCAATTTCCAGAGAACGCCAGATGATGATGA 677  
QY 710 CAACATTTGTGCTCAGTGAAGCAATAGAGTAAAGGTGCTGATCATCGGCAATGCTATG 769  
Db 678 TAAATTTCCGGTGAATTTAGTATGAGATGAGGAGACAGATCATATGCCAATGTTCTG 737  
QY 770 CGAGCCCCAAAACCTTTGGCCCTCTCTTTGAAAATTCGCCATAAATAC 818  
Db 738 CAAGCCTCAGCAACTCAGTGATTACTTCTGAAATTCGGGACAAATTC 786

RESULT 8  
LOCUS AY086590  
DEFINITION Arabidopsis thaliana clone 26006 mRNA, complete sequence.  
ACCESSION AY086590

870 bp mRNA linear PLN 14-APR-2003  
Arabidopsis thaliana clone 26006 mRNA, complete sequence.  
AY086590  
AY086590



```

VERSION AY086590.1 GI:21405300
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 870)
AUTHORS Haas,B.J., Volkovskiy,N., Town,C.D., Troukhan,M., Alexandrov,N.,
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
TITLE Full-length messenger RNA sequences greatly improve genome
annotation
JOURNAL Genome Biol. 3 (6), RESEARCH0029 (2002)
MEDLINE 22089475
PUBMED 12093376
REFERENCE 2 (bases 1 to 870)
AUTHORS Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
TITLE Full-length cDNA from Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 870)
AUTHORS Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
TITLE Direct Submission
JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
Malibu, CA 90265, USA
COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made
available to TIGR and Genbank. The following quality assessment of
this set was done by comparison with known proteins: two percent of
the clones are estimated to be 5'-truncated; less than one percent
are 3'-truncated; approximately two percent represent alternative
splice variants, including unspliced introns and spliced exons; one
percent may contain premature stop codons; five percent may have
frame shifts in a coding region. A sequence is considered to be
5'-truncated if it lacks the translation initiation start (ATG). A
c-terminal end of the encoded protein. Please note that these cDNA
sequences are derived from the Ws or Laer ecotypes and therefore
may contain polymorphisms when compared to sequences from Col-0.
GenSet carried out the library production and sequencing of the
full-length clones. Ceres, Inc. carried out the clustering of the
5' sequences, selection of clones, and sequence assembly.
FEATURES
source
1..870
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="26006"
21..812
/codon_start=1
/product="putative acetone-cyanohydrin lyase"
/protein_id="AA063650.1"
/db_xref="GI:21554666"
/translational="NSEERKQHFVLVHGCHGAWCQWYKVPLEALGHRVTDLAA
SGDITFRSDISTCQYSEPLQMLTSLPNDKQVVLGVHSGFGLSLALAMKFPDKI
SVSVFTAFMPDTHKSPFVEKFPASSMTPEGMWSELETYSGDNLGSLVFTSDPMK
HRLYQLSPVEDLELGLLKPSSLFINELSKMENFSEKGVSPRAYIVCKEDNIIE
DHRQWIMHNPANLVIEETDHPWFCKPQVLSLHLAIDNFS"
CDS
/codon_start=1
/product="putative acetone-cyanohydrin lyase"
/protein_id="AA063650.1"
/db_xref="GI:21554666"
/translational="NSEERKQHFVLVHGCHGAWCQWYKVPLEALGHRVTDLAA
SGDITFRSDISTCQYSEPLQMLTSLPNDKQVVLGVHSGFGLSLALAMKFPDKI
SVSVFTAFMPDTHKSPFVEKFPASSMTPEGMWSELETYSGDNLGSLVFTSDPMK
HRLYQLSPVEDLELGLLKPSSLFINELSKMENFSEKGVSPRAYIVCKEDNIIE
DHRQWIMHNPANLVIEETDHPWFCKPQVLSLHLAIDNFS"
ORIGIN
Query Match 25.3%; Score 273; DB 8; Length 870;
Best Local Similarity 60.3%; Pred. No. 1-2e-57;
Matches 488; Conservative 0; Mismatches 315; Indels 6; Gaps 2;
QY 13 AAAAGAACTACAAGCGCAATAAATTCAAATGAAGGAAGGAAACACTTTGTTTGTAGTAC 72
|||||
Db 1 AAAAAGTACGAAAGGAAAATATGAGTGAGGAGGAAGCAACACTTCGTCTAGTAC 60
|||||
QY 73 ATGGTCATGCCATGGAGGTGGAGTTGGTACAGCTTAAGCCACTGCTAGAGCTGCAG 132
|||||
Db 61 ATGGTGGCTGCCACGGCGCATGGTGGTGTTACAGGTAAAGCTCTCTCGAGGCTTGG 120
|||||
QY 133 GCCATAAGGTTACAGCCCTTGTATTAGCAGCTTCTGCACTGA---TTTGAGAAAAATAG 189
|||||

```

```

121 GCCATCGTGTAAACCGCCTTTAGACCTAGCTGCTTCGGGTATAGACACAACAGGTCAATCA 180
|||||
QY 190 AGGAGCTTCGCACACTTTATGATTATACCTTTGCCATTTGATGGAGTTGATGGAATCTTTT 249
|||||
Db 181 CTGACATTTCTACATGTGTGAACAATATTTCTGAGCCATTGATGAGCTTAATGATCTTCA 240
|||||
QY 250 CAGCAGATGAGAAGGTTATATTAGTGGGGCATAGTCTTGGTGGTATGAATTTGGGACTTG 309
|||||
Db 241 CGAATGATGAGAAGGTTGTACTCGTTGGTCATAGCTTTGAGAGTTTGTAGCCTTAG 300
|||||
QY 310 CTATGGAAGAAGTATCCACAAAAGATCTATGCTGCTGTTTTCTTGGCTGCTTTTCATGCTG 369
|||||
Db 301 CCATGGATAAAGTTTCCCGATAAAATCTCTGTCTCTGTCTCTGCTGCTGCTGCTGCTGCT 360
|||||
QY 370 ATTCTGTTCCACAACCTCTCTCTTTGTTTGGAAACAGTATAATGAGCGGACGCCAGCGAGA 429
|||||
Db 361 ACACCAAAACACTCACCATCGTTGCTGAGAGCAAGATTTGCAAGCAGCATGACACCAAGAG 420
|||||
QY 430 ATTGGTTGGATACTCAGTCTTTTACCATATGTTTCCCTCGAAGAGCCTGACATCCATGT 489
|||||
Db 421 GATGGATGGGCTCTGAGCTCGAGACATATGTT---CAGATAATTCGGCTTGTCTGTCT 477
|||||
QY 490 TTTTGGGCCAAAGTTCTTGGCTCAACAGCTCTACAGCTATGCTCTCTCTGAGGATCTTGG 549
|||||
Db 478 TCTTCAGCACCGACTTCATGAAGCACCGTCTCTACCACTTTCTCTGTGGAGGATCTTGG 537
|||||
QY 550 CATTAGCATCATCTGGTGAGACCAAGCTCTTTGTTTATGGAAGACCTTACGAAGGCCA 609
|||||
Db 538 AGCTTGGATTTGCTTAAAGAGGCTTAGTTCATTTGTTTAAATGAATTTGGAAGATGG 597
|||||
QY 610 AGTATTTCACAGATGAACGGTTTGGATCAGTCAGTGAAGAGAGTTTACATTTGTGTCAC 669
|||||
Db 598 AGAATTTCTGAGAAAGGTTATGGATCTGTTCTCGAGCTTACATTTGTGTCAAAGAGG 657
|||||
QY 670 ATAAAGGCATACAGAAGAAATTCAGCGATGCGAAATTGACAAACATTTGTGTCTACTGA 729
|||||
Db 658 ACAACATTTCTCGGAAGACCATCAACGATGGATGATCCATAATTTATCGGGGAAATTTAG 717
|||||
QY 730 CAATAGAGTTAAAGTGCTGATCAGTCGCAATGCTATGCGAGCCCAAAAACCTTTGCG 789
|||||
Db 718 TCATTGAGATGAAGAGAGCGGATCATATGCCAATGTTTGGCAACCTCAAGTACTAAGTG 777
|||||
QY 790 CTTCTCTCTTGGAAATTTGCCATAAATAC 818
|||||
Db 778 ACCATCTATTTGCAATTCGCTGCAATTTTC 806
|||||
RESULT 9
AF361627
LOCUS AF361627
DEFINITION Arabidopsis thaliana At2g23600/F2686.25 mRNA, complete cds.
ACCESSION AF361627
VERSION AF361627.1 GI:13605602
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 932)
AUTHORS Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Banh,J., Bowser,L.,
Carninci,P., Chung,M.K., Goldsmith,A.D., Hayashizaki,Y., Ishida,J.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M.,
Lin,J., Liu,S.X., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
Phan,P.K., Quach,H.L., Sakano,H., Sakurai,T., Satou,M., Seki,M.,
Southwick,A., Toriumi,M., Yamada,K., Yu,G., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
TITLE Arabidopsis cDNA clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 932)
AUTHORS Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Banh,J., Bowser,L.,
Carninci,P., Chung,M.K., Goldsmith,A.D., Hayashizaki,Y., Ishida,J.,

```





```
Qy 115 CACTGCTAGAGCTGCAGGCCATGAAGTTTACAGCCCTTGATTTAGCAGCTTCTGGCACTG 174
Db 83 CTCTTCTCGAGGCTTTGGGCCATCGTGTAAACCGCTTAGACCTAGCTGCTTCGCGTATAG 142
Qy 175 A---TTTGAGAAAATAGAGAGCTTCGCACACTTTATGATTTATCTATCTTGGCATTGATGG 231
Db 143 ACACAACAGGTCAATCACTGACATTTCTATCATGTGTGAACAATATCTGAGCCCAATTGATGC 202
Qy 232 AGTTGATGGAACTCTCTTTCAGCAGATGAGAAGGTTATATTAGTGGGCGCATAGTCTTGGTG 291
Db 203 AGCTAATGACTTCAATTCGCCGAATGATGAGAAGGTTGTACTGTTGGTCATAGCTTTGGAG 262
Qy 292 GTATGAATTTGGGACTTGTCTATGAAAGATATCAACAAGAGCTATGCTGCTGTTTCT 351
Db 263 GTTTGAGTTTAGCTTAGCCATGATGAATGTTTCCGATAAAATCTCTGCTCTGCTCTG 322
Qy 352 TGCTGCTTTCATGCTCTGATTTCTGTTACAACTCTCTTGTGTTTGGACAGTAAATG 411
Db 323 TGACTGCAATTCATGCCCGCACCAAAACATCATCATCGTTCTGTCGAGGAAAGTTTGCAA 382
Qy 412 AGCGGACGCCAGCCGAGAAATTTGGTGGATCTACTCAGTTTATACCATATGGTTCCCTGAAG 471
Db 383 GCAGCATGACACCAAGAGATGATGGCTCTGAGCTCGAGACATATGTT---CAGATA 439
Qy 472 AGCCACTGACATCCATGTTTTTTGGCCAAAGTTCTTGGCTCAAAAGCTCTACCACTAT 531
Db 440 ATTCGCGCTTGTCTGTGTTCTTCAGCACCGACTTCATGAAGCACCGCTCTCTACCACTTT 499
Qy 532 GCTCTCTGAGGATCTTGCAATAGCATCATCTGTTGGTGAAGCAAGCTCTTTGTTATGG 591
Db 500 CTCCTGTGGAGGATCTTGAGCTTTGGATGTTCTTAAAGAGCGCTAGTTCATTGTTTATTA 559
Qy 592 AAGACCTATCGAAGGCCAAGTATTTCCACAGATCAACGTTTGGATCAGTGAAGAGGTTT 651
Db 560 ATGAATATCGAAGATGAGAACTTTTCTGAGAAAGGTTATGATCTGTTCTCGAGCTT 619
Qy 652 ACATTGTTGCTCACTGAGGATAAAGGCATACCAAGAAATTTCCAGCGATGCGCAATTGACA 711
Db 620 ACATTGTTGCAAGAGGACAACTATTCGGAAGACCATCAACCATGATGATCCATA 679
Qy 712 ACATTGTTGCTCACTGAGCAATAGATTAAGGTTGCTGATCATGATGCAATGCTATGCG 771
Db 680 ATTTATCGCGGAATTTAGTGAATGAGATGAGAGAGACTGATCATATATGCGCAATGTTTGA 739
Qy 772 AGCCCCAAAACCTTTGCGCTCTCTCTTGGAAATTCGCCATAAATAC 818
Db 740 AACCTCACTAAGTGACCATCTATTTGGCAATCGCTGACAAATTC 786
```

```
RESULT 11
AY056068 792 bp mRNA linear PLN 30-SEP-2001
LOCUS Arabidopsis thaliana At2g23600/F26B6.25 mRNA, complete cds.
DEFINITION AY056068
ACCESSION AY056068.1 GI:15810084
VERSION FLI CDNA.
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 792)
AUTHORS Koesema, E., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Shinn, P.,
Ban, J., Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D.,
Hayashizaki, Y., Iehida, J., Jiang, P.X., Jones, T., Kamiya, A.,
Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X.,
Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J.,
Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A.,
Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,
Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
Arabisidopsis ORF clones
Unpublished
REFERENCE 2 (bases 1 to 792)
```

## AUTHORS

Koesema, E., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Shinn, P.,  
Ban, J., Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D.,  
Hayashizaki, Y., Iehida, J., Jiang, P.X., Jones, T., Kamiya, A.,  
Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X.,  
Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J.,  
Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A.,  
Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,  
Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

## Direct Submission

Submitted (11-SEP-2001) Salk Institute Genomic Analysis Laboratory  
(SIGNAL), Plant Biology Laboratory, The Salk Institute for  
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,  
USA

## COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN  
Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Iehida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the  
sequencing and annotation of the RAFL cDNAs: Koesema, E., Chen, H.,  
Cheuk, R., Kim, C.J., Meyers, M.C., Shinn, P., Ban, J., Bowser, L.,  
Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G.,  
Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,  
Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A.,  
Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,  
Davis, R.W., Theologis, A., and Ecker, J.R.

Koesema, E. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally  
to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)  
contributed equally to this work as PIs.

## FEATURES

## Source

1..792  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="U12118"  
/ecotype="Columbia"  
/note="This clone is in PUNI 51"

## CDS

1..792  
/note="putative acetone-cyanohydrin lyase"  
/codon\_start=1  
/product="At2g23600/F26B6.25"  
/protein\_id="AAL06968.1"  
/db\_xref="GI:15810085"  
/translation="MSEERKQHFVLVHGCHGAWCYKVPLEALGHRVATLDLAA  
SGIDTTSITDISTCEQYSEFLMQLMPLNDEKVLVGHSGLSLALAMDKFPDK  
SVSVFTAFMPDTHSPFVEKFASSMTPEGMGSELETGSDNSGLSVFSTDFMK  
HRLYOLSPVEDLELGLKRPSSLFPINELSKMENPEKGYGVPRAIVCKEDNIIE  
DQRWMIHNPANLVIEETDHPMFCKPQLSLDHLAIADNFC"

## ORIGIN

Query Match 25.1%; Score 271; DB 8; Length 792;  
Best Local Similarity 61.4%; Pred. No. 3.7e-57;  
Matches 471; Conservative 0; Mismatches 290; Indels 6; Gaps 2;  
Qy 55 AACACTTTGTTTGTAGTACATGGTGCATGCCATGGAGTTGGAGTTGGTACAGCTAAAGC 114  
Db 23 AACACTTCTGTGTAGTACATGGTGCATGCCATGGAGTTGGTACAGCTAAAGC 82  
Qy 115 CACTCTAGAGCTGCAGGCCATAGGTTTACAGCCCTTGTATTAGCAGCTTCTTGGCAGCTG 174  
Db 83 CTCCTCTCGAGCTTTGGGCCATCGTGTAAACCGCTTAGACCTAGCTGCTTCGCGTATAG 142  
Qy 175 A---TTTGAGAAAATAGAGGAGCTTCGCACACTTTATGATTTATCTTTGCCATTGATGG 231  
Db 143 ACACAACAGGTCATCACTGACATTTCTACATGTGAACAATATTTCTGAGCCATTGATGC 202  
Qy 232 AGTTGATGGAATCTCTTTTCAGCAGATGAGAAGGTTTATATTAGTGGGCGCATATCTTGGTG 291  
Db 203 AGCTAATGACTTTCATTTGCCGAATGATGAGAAGGTTGTACTCGTTGGTCATAGCTTTGGAG 262

TITLE  
JOURNAL  
REFERENCE

```
QY 292 GTATGAATTTGGCACTTGGCTATGGAAAAAGATATCCACAAAGATCTATGCTGCTGTTTTCT 351
DB 263 GTTTGAGTTTGGCTTTAGCCATGGATTAAGTTTCCCGATAAAATCTCTGCTCTGCTTGG 322
QY 352 TGGCTGCTTTTCATGCTGCTGATTTCTGTTACAACTCTCTCTCTTTGTTTGGAAACAGTATAATG 411
DB 323 TGACTGCATTCATGCCCCGACACCAACACTCCACCATCGTTCTCGAGGAAAGTTTGCA 382
QY 412 AGCGAGCCGACCGAGAAATGGTTGGATPACTCAGTTTTTACCATATGGTTCCCTCGAAG 471
DB 383 GCAGCATGACACAGAAAGATGGAATGGGCTCTGAGCTCGAGACATATGGTT--CAGATA 439
QY 472 AGCCACTGACATCCATGTTTTTTGGCCCAAGTTCTTGGCTCACAAGCTCTACCGACTAT 531
DB 440 ATTCGGGCTTGTGTGTTCTTTAGCAGCCGACTTCATGAAGCACCCTCTCTACCAACTTT 499
QY 532 GCTCTCTCGAGATCTTGATAGCATCATCTGTTGGTAGACCAAGCTCTTTGTTTATGG 591
DB 500 CTCCTGTGAGGATCTTGAGTTGGATTCCTTAAGAGGCTAGTTCATTTGTTTATTA 559
QY 592 AAGACCTATCGAAGGCCAAAGTATTTTCCAGATCAACGGTTTGGATCAGTGAAGAGAGTTT 651
DB 560 ATGAATTTATCGAAGATGGAGAACTTTCTGAGAAAGGGTATGGATCTGTTCTCGAGCTT 619
QY 652 ACATTTGTGCTAGGATTAAGGATACCAAGATACCAAGAAATTCAGCGATGGCAATTGACA 711
DB 620 ACATTTGTGTCGAAGAGGACCAATTTATCTCGGAAGACCATCAACGATGGATGCCATA 679
QY 712 ACATTTGTGCTAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 771
DB 680 ATTATCCGGGGAATTTAGTATGATGATGAGATGAGAGACTGATCATATGCCAATGTTTGA 739
QY 772 AGCCCCAAAACCTTTGCGGCTCTCTCTTGGAAATTTGCCATAAATAC 818
DB 740 AACCTCACTACTAAGTGACCATCTATTGGCAATGCTGACAAATTC 786

RESULT 12
CQ806104 792 bp DNA linear PAT 10-MAY-2004
LOCUS Sequence 2515 from Patent WO2004035798.
DEFINITION CQ806104
ACCESSION CQ806104.1 GI:47111686
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1
REFERENCE Inze,D., de Veylder,L. and Vlieghe,K.
AUTHORS Identification of novel e2f target genes and use thereof
TITLE Patent: WO 2004035798-A 2515 29-APR-2004;
JOURNAL CropDesign N.V. (BE)
FEATURES Location/Qualifiers
source 1..792
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match 24.9%; Score 269; DB 6; Length 792;
Best Local Similarity 61.1%; Pred. No. 1.2e-56;
Matches 472; Conservative 0; Mismatches 295; Indels 6; Gaps 2;
QY 55 AACACTTTCTTTAGTACATGGTGCATGCCATGGAGTTGGAGTTGGTACAGCTAAAGC 114
DB 23 AACACGTCGTTCTAGTACATGGTGTGCTGCGATGGCCCTGGTGTACAGGTTAAGC 82
QY 115 CACTGCTAGAAAGCTCAGGCGCAATAGGTTACAGCCCTTGATTTTAGCAGCTTCTGGCAGTG 174
DB 83 CGCAGCTCGAGGCTTCTGGCCACCGGTAAACCGCGTAGATCTAGTGTCCGGTATAG 142
```

```
QY 175 ATTTG--AGAAAATAGAGGAGCTTCGCACACTTTATGATTATATCTTTGCCATTGATGG 231
DB 143 ACATGACCAGGTCAATCAAGATATATCCATGCGGAACAATATCTCAGAGCAATTGATGC 202
QY 232 AGTTGATGGAATCTCTTTTACGACAGATGAGAAGGTTATATTAGTGGGGCATAGTCTTGTTG 291
DB 203 AGCTAATGACCTCACTACCAATGATGAGAAGGTTGTGCTGTTGGTCATAGCTTAGGAG 262
QY 292 GTATGAATTTGGGACTTGTCTATGGAATAAGATATCCAAAAAGATCTATGCTGCTGTTTTCT 351
DB 263 GTTTGAGTTTAGCTATGGCCATGGATATGTTTCCGACCAAAATCTCTGTTCTGTCTTTG 322
QY 352 TGGCTGCTTTTCATGCTGATTTCTGTTCAACTCTCTCTCTTTGTTTGGAAACAGTATAATG 411
DB 323 TGACTGCATGATGCGACAGACCAACACTCACCATCTCTGATGGGATAGCTTAAGAA 382
QY 412 AGCGAGCCGACCGAGAAATGGTTGGATATCTCAGTATTTTACCATATGTTTCCCTCGAAG 471
DB 383 AAGAAACTTTCAGAGAGGAATGGTTAGACACCGTGT--ACGAGCGAGAAACCTGATT 439
QY 472 AGCCACTGACATCCATGTTTTTTGGCCCAAGTTCTTGGCTCACAAGCTCTACCGACTAT 531
DB 440 TTCTAGCGAGTTTGGATTTTGGACCAAGAAATTCATGCCCAAGAACTTGTATCAGTTGT 499
QY 532 GCTCTCTCGAGGATCTTGATAGCATCATCTGTTGGTAGACCAAGCTCTTTGTTTATGG 591
DB 500 CTCAGTCCAAAGATCTTGAATGGCGAAATGTTGGTGAGGGCAACCACTGATTGAAGA 559
QY 592 AAGACCTATCGAAGGCCAAAGTATTTTCCAGATCAACGGTTTGGATCAGTGAAGAGAGTTT 651
DB 560 AAGATATGCGAGAGAGAGAGCTTCTAGTGAAGAGGATACGGATCCGTACACGATAT 619
QY 652 ACATTTGTGCTAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 711
DB 620 TTATTTGATGCGAAAGGATCTTGTGTCAACCGAAGATTAACGAGGATCGATGATCAGCA 679
QY 712 ACATTTGTGCTAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 771
DB 680 ACTTTCCCCCAAGAGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 739
QY 772 AGCCCCAAAACCTTTGCGGCTCTCTCTTGGAAATTTGCCATAAATAC 824
DB 740 AGCCTCAACAACTATGCTCTCTCTTGGAGATTGCAAAATAAATATGCCATAA 792

RESULT 13
AX505522 792 bp DNA linear PAT 27-SEP-2002
LOCUS Sequence 217 from Patent WO0216655.
DEFINITION AX505522
ACCESSION AX505522.1 GI:23386759
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1
REFERENCE Harper,J.F., Krens,J., Wang,X. and Zhu,T.
AUTHORS Stress-regulated genes of plants, transgenic plants containing
TITLE same, and methods of use
JOURNAL Patent: WO 0216655-A 217 28-FEB-2002;
FEATURES The Scripps Research Institute (US) ; Syngenta Participations AG (CH)
Location/Qualifiers
source 1..792
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match 24.9%; Score 269; DB 6; Length 792;
Best Local Similarity 61.1%; Pred. No. 1.2e-56;
```

| RESULT 14  | BT015031  | 792 bp      | linear | PLN 10-JUL-2004 |
|------------|---|-------------|--------|-----------------|
| LOCUS      | BT015031  |             | mrna   |                 |
| DEFINITION | Arabidopsis thaliana At2g23610 mRNA, complete cds.  |             |        |                 |
| ACCESSION  | BT015031  |             |        |                 |
| VERSION    | BT015031.1  | GI:50198957 |        |                 |
| KEYWORDS   | FLI CDNA.   |             |        |                 |
| SOURCE     | Arabidopsis thaliana (thale cress)  |             |        |                 |
| ORGANISM   | Arabidopsis thaliana  |             |        |                 |
|            | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots, rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi. |             |        |                 |
| REFERENCE  | 1 (bases 1 to 792)  |             |        |                 |
| AUTHORS    | Shinn, P., Chen, H., Cheuk, R., Kim, C.J. and Ecker, J.R.   |             |        |                 |
| TITLE      | Arabidopsis ORF clones  |             |        |                 |
| JOURNAL    | Unpublished   |             |        |                 |
| REFERENCE  | 2 (bases 1 to 792)  |             |        |                 |
| AUTHORS    | Shinn, P., Chen, H., Cheuk, R., Kim, C.J. and Ecker, J.R.   |             |        |                 |

```

Db      620 TTATTGTATGCGGAAGGATCTTGTCACCCGAGATTAACACGCGATGATCAGCA 679
QY      712 ACATTGGTGTCTACTGAAGCAATAGAGATTAAAGGTGCTGATCACATGCGAATGCTATCGG 771
Db      680 ACTTTCCCCAAAGAAGTAATGGAGATCAAGAAGCGCAGATCATATGCGCAATGTTCTCCA 739
QY      772 AGCCCCAAAACATTTGGCGCTCTCTCTGGAAATTTGCCCAATAAATACAACATGA 824
Db      740 AGCCTCAACAACATATGTGCTCTCTCTGGAGATTGCAAAATAAATATGCGCTAA 792

RESULT 15
LOCUS   BT014901
DEFINITION Arabidopsis thaliana At2g23610 gene, complete cds.
ACCESSION BT014901
VERSION   BT014901.1 GI:48596978
KEYWORDS FLI CDNA.
SOURCE   Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 878)
          Kim,C.J., Chen,H., Cheuk,R., Shinn,P. and Ecker,J.R.
          Arabidopsis cDNA clones
          Unpublished
REFERENCE 2 (bases 1 to 878)
          Kim,C.J., Chen,H., Cheuk,R., Shinn,P. and Ecker,J.R.
          Direct Submission
          Submitted (11-JUN-2004) Salk Institute Genomic Analysis Laboratory
          (SIGNAL), Plant Biology Laboratory, The Salk Institute for
          Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
          USA

FEATURES             Location/Qualifiers
     source            1..878
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosomes="2"
                     /clone="S67063"
                     /ecotype="Columbia"
                     /note="this clone is in pUNI 51"
     CDS               8..799
                     /note="unknown protein"
                     /codon_start=1
                     /product="At2g23610"
                     /protein_id="AT46030.1"
                     /db_xref="GI:48596979"
                     /translation="MSEERKQHVVLVHGACHGAWCWYKVKQLRSGHRVAVDLAA
                     SGIDMTRSTIDTCEQYSEPLQMLSLPDDERKVLVGHSLGSLSLAMVMDMPFKI
                     SVSVFTAMPPTKISPSFVWDLRKRTSREWLDTFTSEKPDFSEFWIFGPPEFMA
                     KNYQLSPVQDLLELAKMLVRANPLIKDMAERRSFSEGYGSVTRI FIVCGKDLVSPE
                     DYQRSMISNFPPEVMEIKDADMPFSPQQLCALLLEIANKYA"

ORIGIN
Query Match      24.9%; Score 269; DB 8; Length 878;
Best Local Similarity 61.1%; Pred. No. 1.2e-56;
Matches 472; Conservative 0; Mismatches 295; Indels 6; Gaps 2;
QY      55 AACACTTTGTTTAGTACATGGTGCATGCCATGGAGGTTGGAGTTGGTACAAAGCTAAAGC 114
Db      30 AACACGTCGTTCTAGTACATGGTGTCTTGGCATGGCGCCTGGTGTACAAAGGTTAAGC 89
QY      115 CACTGTCTAGAGCTGCAGGCATTAAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCACTG 174
Db      90 CGCAGCTCGAGGCTTCTGGCCACCGCGTACCGCCGTAGATCTAGCTGCGTATAG 149
QY      175 ATTG---AGAAAAATAGAGGAGCTTCGCACACTTTATGATTATACTTTGCCATTGATGG 231
Db      150 ACATGACCAAGTCAATCACAGATATATCCATGCGAACAATACTCAGAGCCATTGATGC 209
QY      232 AGTTGATGGATCTCTTTAGCAGATGAGAAGGTTATATAGTGGGGCATAGTCTTGGTG 291
```

```

Db      210 AGCTAATGACCTCACTACCAGATGATGAGAAGGTTGTGCTTGTGGTCATAGCTTAGGAG 269
QY      292 GTATGAATTTGGGACTCTGCTATGGAAGATATCCAAAAGATCTATGCTGCTGCTTTCT 351
Db      270 GTTTGAGTTTAGCTATGGCCATGATGTTTCCGACCAAAATCTCTGTTTCTGCTTTG 329
QY      352 TGGCTGCTTTTCATGCTGATTTCTGTTCAACTCTCTCTTGTGTTTGGAAACAGTATATG 411
Db      330 TGACTGCTATGATGCCAGACACCAAACTCACCATCTCTCGTATGGATTAAGCTAAGAA 389
QY      412 AGCGGACGCCAGCCGAGAAATGGTTGGATCTCAGTTTTCACCATATGTTTCCCTTGAAG 471
Db      390 AAGAAACTTTCAGAGAGGAATGGTTAGACACCGTGT---ACGAGCGAGAAACCTGATT 446
QY      472 AGCCACTGACATCCATGTTTTTTGGCCCAAAGTCTTTGGGCTCACAAGCTCTACCACTAT 531
Db      447 TTCCTAGCGAGTTTGGATTTTGGACCCAGAAATTCATGGCCCAAGAACTTGATCAGTTGT 506
QY      532 GCTCTCCTGAGGATCTTGCATTAGCATCATCTTGGTGAACCGTTTGGATCAGTGAAGAGATT 651
Db      507 CTCCAGTCCAAAGATCTTGAATTTGGCGAAATGTTGGTGAGGGCAACCCATTTGATTAGA 566
QY      592 AAGACCTATCGAAGGCCCAAGTATTTTCACAGATGAACCGTTTGGATCAGTGAAGAGATT 651
Db      567 AAGATATGCCAGAGAGAGAAGCTTTCAGTGAGGAAGGATACGGATCCGTTACACGTATAT 626
QY      652 ACATTGTGTGCACTGAGGATAAAGCATACAGAAAGAAATTCAGCGATGGCAAAATTGACA 711
Db      627 TTATTGTATGCGGAAAAGGATCTTGTGTCAACCGAAGATTACCGCGATCGATGATCAGCA 686
QY      712 ACATTGGTGTCACTGAAGCAATAGATTAAAGGTCTGATGATCAGATGCGCAATGCTATCGG 771
Db      687 ACTTTCCCCCAAAAAGAGTAATGGAGATCAAAAGACGCAGATCATATGCCAATGTTCTCCA 746
QY      772 AGCCCCAAAAACTTTGCGCCTCTCTCTTGGAAATTTGCCCATATAAATAFCAACTGA 824
Db      747 AGCCTCAACAACATATGTGCTCTTCTCTGGAGATTGCAAAATAATATGCTTAA 799
```

Search completed: September 12, 2005, 04:45:50  
Job time : 4689.27 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: September 11, 2005, 22:50:50 ; Search time 591.255 Seconds  
(without alignments)  
10803.121 Million cell updates/sec

Title: US-10-780-002-1  
Perfect score: 1079  
Sequence: 1 acgcggggaagaagaaagaaa.....gaaaaaaaaaaaaaaaaaaaaa 1079

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues  
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID       | Description        |
|------------|-------|-------------|--------|-------------|--------------------|
| 1          | 1079  | 100.0       | 1079   | 8 ADA26498  | Ada26498 Tobacco s |
| 2          | 295.2 | 27.4        | 918    | 8 ADA26530  | Ada26530 Arabidops |
| 3          | 292.2 | 27.1        | 792    | 6 ABZ13515  | Abz13515 Arabidops |
| 4          | 292.2 | 27.1        | 816    | 3 AAC42344  | Aac42344 Arabidops |
| 5          | 276.2 | 25.6        | 950    | 8 ADA26531  | Ada26531 Arabidops |
| 6          | 273   | 25.3        | 870    | 3 AAC34411  | Aac34411 Arabidops |
| 7          | 272.6 | 25.3        | 864    | 3 AAC47519  | Aac47519 Arabidops |
| 8          | 271.6 | 25.2        | 1152   | 8 ADA26532  | Ada26532 Arabidops |
| 9          | 271   | 25.1        | 792    | 6 ABZ13746  | Abz13746 Arabidops |
| 10         | 269   | 24.9        | 792    | 3 AAC43133  | Aac43133 Arabidops |
| 11         | 269   | 24.9        | 792    | 6 ABZ12412  | Abz12412 Arabidops |
| 12         | 269   | 24.9        | 792    | 12 ADN74620 | Adn74620 Thale cre |
| 13         | 256   | 23.7        | 1109   | 8 ADA26533  | Ada26533 Arabidops |
| 14         | 252.6 | 23.4        | 1046   | 8 ADA26535  | Ada26535 Arabidops |
| 15         | 243.2 | 22.5        | 928    | 8 ADA26538  | Ada26538 Arabidops |
| 16         | 242.8 | 22.5        | 823    | 3 AAC42529  | Aac42529 Arabidops |
| 17         | 240.8 | 22.3        | 834    | 8 ADA26537  | Ada26537 Arabidops |
| 18         | 239.8 | 22.2        | 771    | 6 ABZ13473  | Abz13473 Arabidops |
| 19         | 239.4 | 22.2        | 770    | 8 ADA67827  | Ada67827 Arabidops |
| 20         | 234.2 | 21.7        | 1056   | 3 AAC47895  | Aac47895 Arabidops |

|    |       |      |      |             |                     |
|----|-------|------|------|-------------|---------------------|
| 21 | 232.6 | 21.6 | 909  | 3 AAC47800  | Aac47800 Arabidops  |
| 22 | 232.6 | 21.6 | 984  | 8 ADA26534  | Ada26534 Arabidops  |
| 23 | 224.8 | 20.8 | 783  | 8 ADA26536  | Ada26536 Arabidops  |
| 24 | 215   | 19.9 | 903  | 3 AAC34141  | Aac34141 Arabidops  |
| 25 | 215   | 19.9 | 908  | 8 ADA26539  | Ada26539 Arabidops  |
| 26 | 212.8 | 19.7 | 828  | 3 AAC41733  | Aac41733 Arabidops  |
| 27 | 186   | 17.2 | 1091 | 2 AAT36351  | Aat36351 Hevea bra  |
| 28 | 184.2 | 17.1 | 817  | 2 AAT86824  | Aat86824 Hevea bra  |
| 29 | 169   | 15.7 | 777  | 12 ADQ16305 | Adq16305 Nucleotid  |
| 30 | 168   | 15.6 | 795  | 8 ADA70991  | Ada70991 Rice gene  |
| 31 | 156.6 | 14.5 | 777  | 12 ADQ16306 | Adq16306 Nucleotid  |
| 32 | 149   | 13.8 | 422  | 3 AAC43356  | Aac43356 Arabidops  |
| 33 | 145.8 | 13.5 | 1325 | 3 AAC47916  | Aac47916 Arabidops  |
| 34 | 143.4 | 13.3 | 984  | 3 AAC49239  | Aac49239 Arabidops  |
| 35 | 126.2 | 11.7 | 483  | 10 ADE82083 | Ades82083 Arabidops |
| 36 | 125.2 | 11.6 | 1155 | 10 ART23286 | Art23286 Seed deve  |
| 37 | 111.4 | 10.3 | 121  | 12 ADG00027 | Adg00027 Nicotiana  |
| 38 | 108.2 | 10.0 | 1550 | 8 ADA26540  | Ada26540 Arabidops  |
| 39 | 107.4 | 10.0 | 541  | 10 ABX57306 | Abx57306 Arabidops  |
| 40 | 105.4 | 9.8  | 1371 | 8 ADA26544  | Ada26544 Arabidops  |
| 41 | 103.2 | 9.6  | 1815 | 3 AAC44497  | Aac44497 Zea maye   |
| 42 | 102   | 9.5  | 860  | 8 ADA711305 | Ada711305 Rice gene |
| 43 | 95.6  | 8.9  | 1665 | 8 ADA26542  | Ada26542 Arabidops  |
| 44 | 94.8  | 8.8  | 783  | 8 ADA70766  | Ada70766 Rice gene  |
| 45 | 92.6  | 8.6  | 771  | 3 AAC43020  | Aac43020 Arabidops  |

ALIGNMENTS

RESULT 1  
ADA26498  
ID ADA26498 standard; cDNA; 1079 BP.  
XX  
AC ADA26498;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Tobacco salicylic acid binding protein 2 (SABP2) gene.  
XX  
KW salicylic acid-binding protein; SABP2; disease resistance; plant;  
modulator; tobacco; ds; gene.  
XX  
OS Nicotiana tabacum.  
XX  
FH Key Location/Qualifiers  
CDS 42..824  
FT /\*tag= a  
FT /product= "salicylic acid binding protein SABP2"  
XX  
PN WO2003016551-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 16-AUG-2002; 2002WO-US026312.  
XX  
PR 16-AUG-2001; 2001US-0312863P.  
XX  
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.  
XX  
PI Klessig DF, Kumar D;  
DR WPI; 2003-278578/27.  
DR P-PSDB; ADA26499.  
XX  
PT New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for  
producing transgenic plants with increased resistance to disease, or for  
screening SABP2 modulators that confer enhanced resistance of plants to  
disease.  
XX  
PS Claim 1; Fig 4; 99pp; English.  
CC The invention relates to a novel isolated 1079 bp salicylic acid-binding

CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid  
CC protein, sequences that specifically hybridize with the nucleic acid, the  
CC complement of the nucleic acid or a natural allelic variant of the  
CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for  
CC conferring increased resistance to disease in plants, or for screening  
CC modulators of SABP2, which confer increased or enhanced resistance of  
CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also  
CC useful for producing transgenic plants with increased resistance to  
CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful  
CC for studying the molecular mechanisms responsible for the modulation of  
CC salicylic acid-mediated disease resistance in plants. This sequence  
CC represents the gene encoding the SABP2 protein from tobacco plants.  
XX  
SQ Sequence 1079 BP; 307 A; 204 C; 234 G; 334 T; 0 U; 0 Other;

Query Match 100.0%; Score 1079; DB 8; Length 1079;  
Best Local Similarity 100.0%; Pred. No. 1e-279;  
Matches 1079; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGCGGGGAAAGAAAGAACTAAACAGGCATATAATTCAAATGAAGGAGGAAACACT 60  
DB 1 ACGCGGGGAAAGAAAGAACTAAACAGGCATATAATTCAAATGAAGGAGGAGGAAACACT 60

QY 61 TTGTTTACATGTCATGCCATGCCATGGAGGTTGGAGTTGGTACAGCTAAGGCCACTGC 120  
DB 61 TTGTTTACATGTCATGCCATGCCATGGAGGTTGGAGTTGGTACAGCTAAGGCCACTGC 120

QY 121 TAGAGCTCAGGCCATAAGGTTTACAGCCCTTGATTTAGCAGCTTCTGGCACTGATTTGA 180  
DB 121 TAGAGCTCAGGCCATAAGGTTTACAGCCCTTGATTTAGCAGCTTCTGGCACTGATTTGA 180

QY 181 GAAATAAGAGGAGCTTCCACACTTTATGATATACCTTTGCAATGATGGAGTTGATGG 240  
DB 181 GAAATAAGAGGAGCTTCCACACTTTATGATATACCTTTGCAATGATGGAGTTGATGG 240

QY 241 AATCTCTTTACAGAGTAGAGGTTATATTAGTGGGCATAGTCTTGGTGGTATGAAAT 300  
DB 241 AATCTCTTTACAGAGTAGAGGTTATATTAGTGGGCATAGTCTTGGTGGTATGAAAT 300

QY 301 TGGGACTTGTATGGAAAAGTATCCAAAAAGATCTATGCTGTCTTTCTTGGCTGCTT 360  
DB 301 TGGGACTTGTATGGAAAAGTATCCAAAAAGATCTATGCTGTCTTTCTTGGCTGCTT 360

QY 361 TCATGCTGATTCGTGTCACACTCTCTTTGTTTGGAAACAGTATATAGCGGACGC 420  
DB 361 TCATGCTGATTCGTGTCACACTCTCTTTGTTTGGAAACAGTATATAGCGGACGC 420

QY 421 CAGCCGAGAAATGGTTGGATACCTCAGTCTTTTACCATATGTTCCCTGAAGGCCACTGA 480  
DB 421 CAGCCGAGAAATGGTTGGATACCTCAGTCTTTTACCATATGTTCCCTGAAGGCCACTGA 480

QY 481 CATCCATGTTTTTGGCCCCAAGTTCTTGGCTCACAAGCTCTACCAAGCTATGCTCTCCTG 540  
DB 481 CATCCATGTTTTTGGCCCCAAGTTCTTGGCTCACAAGCTCTACCAAGCTATGCTCTCCTG 540

QY 541 AGGATCTGCATPAGCATCATATGGTGAGACCAAGCTCTTTGTTTATGGAAGCCTAT 600  
DB 541 AGGATCTGCATPAGCATCATATGGTGAGACCAAGCTCTTTGTTTATGGAAGCCTAT 600

QY 601 CGAAGCCCAAGTATTTTACAGATGACGGTTTGGATCAGTGAAGAGGTTTACATTTGCT 660  
DB 601 CGAAGCCCAAGTATTTTACAGATGACGGTTTGGATCAGTGAAGAGGTTTACATTTGCT 660

QY 661 GCACTGAGGATAAGGCATACCAAGAAATTCAGCGATGGCAATTTGACCAATTTGGTG 720  
DB 661 GCACTGAGGATAAGGCATACCAAGAAATTCAGCGATGGCAATTTGACCAATTTGGTG 720

QY 721 TCATGAAGCAATPAGAGATTAAGGTGCTGATCAGATGCAATGCAATGCGAGCCCAAA 780  
DB 721 TCATGAAGCAATPAGAGATTAAGGTGCTGATCAGATGCAATGCGAGCCCAAA 780

QY 781 AACTTTGGCCTCTCTCTTGGAAATTTGCCATATAATACAACTGATCTCTACATATGCT 840

DB 781 AACTTTGGCCTCTCTCTTGGAAATTTGCCATAAATACAACTGATCTCTACATATGCT 840

QY 841 TCGTCTCATGTCAAGATTTTCAGTGCAATGCTGTAATTTTCTATTTTTCGACCGGCGC 900

DB 841 TCGTCTCATGTCAAGATTTTCAGTGCAATGCTGTAATTTTCTATTTTTCGACCGGCGC 900

QY 901 ATACTGTCTTTGGCTATTTTAAGGATGCAATTTTCACTCTCTAGTGTGGAAGGCT 960

DB 901 ATACTGTCTTTGGCTATTTTAAGGATGCAATTTTCACTCTCTAGTGTGGAAGGCT 960

QY 961 TCCACATAAGGATTTGCTTCTGTTTCTCCATTCAAGTGTGTTATGTTGAGATACTTAAAC 1020

DB 961 TCCACATAAGGATTTGCTTCTGTTTCTCCATTCAAGTGTGTTATGTTGAGATACTTAAAC 1020

QY 1021 CGTATCAATCTTGTAATGAAACCTTCTTCTTCTTTTGAIAAAAAAAAAAAAA 1079

DB 1021 CGTATCAATCTTGTAATGAAACCTTCTTCTTCTTCTTTTGAIAAAAAAAAAAAAA 1079

RESULT 2  
ADA26530  
ID ADA26530 standard; DNA; 918 BP.  
XX  
AC ADA26530;  
XX AC  
XX 20-NOV-2003 (first entry)  
XX  
DE Arabidopsis thaliana AtSB2L1 gene homologous to tobacco SABP2 gene.  
XX  
KW salicylic acid-binding protein; SABP2; disease resistance; plant;  
KW modulator; tobacco; ds; gene.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO2003016551-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 16-AUG-2002; 2002WO-US026312.  
XX  
PR 16-AUG-2001; 2001US-0312863P.  
XX (BOYC-) BOYCE THOMPSON INST PLANT RES INC.  
XX  
PI Klessig DF, Kumar D;  
XX  
DR WPI; 2003-278578/27.  
XX  
PT New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for  
PT producing transgenic plants with increased resistance to disease, or for  
PT screening SABP2 modulators that confer enhanced resistance of plants to  
PT disease.  
XX  
PS Claim 65; Page; 98pp; English.  
XX  
CC The invention relates to a novel isolated 1079 bp salicylic acid-binding  
CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid  
CC protein, sequences that specifically hybridize with the nucleic acid, the  
CC complement of the nucleic acid or a natural allelic variant of the  
CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for  
CC conferring increased resistance to disease in plants, or for screening  
CC modulators of SABP2, which confer increased or enhanced resistance of  
CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also  
CC useful for producing transgenic plants with increased resistance to  
CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful  
CC for studying the molecular mechanisms responsible for the modulation of  
CC salicylic acid-mediated disease resistance in plants. This sequence  
CC represents the gene encoding the Arabidopsis thaliana SABP2 AtSB2L1 which  
CC has homology to the gene (ADA26498) encoding the tobacco SABP2 protein.  
CC (Note: this sequence is not given in the specification but is derived  
CC from the Genbank accession number given by the inventors).  
XX  
SQ Sequence 918 BP; 266 A; 175 C; 221 G; 256 T; 0 U; 0 Other;



```
Query Match      27.4%; Score 295.2; DB 8; Length 918;
Best Local Similarity 62.7%; Pred. No. 6.9e-69;
Matches 494; Conservative 0; Mismatches 288; Indels 6; Gaps 2;

QY 53 AAAACACTTTGTTTGTAGTACATCGGTGCATGCCATGAGGTTGGAGTTGGTACAAAGCTAAA 112
DB 64 ACAACATTTTGTACTAGTACATCGTTCTGTCATGCGGTGGTGTGCTGTGTTACAAAGTTAA 123

QY 113 GCCACTGTAGAGCTGCAGGCCATAGGTTTACAGCCCTTGATTTAGCAGCTTCTGGCAC 172
DB 124 GCCGCTGTAGAGCGGTGGGCCACCGCGTAACCTGTGTGAGCTTAGCTGCGCTCCGGAAT 183

QY 173 TGA---TTTGAGAAAAATAGAGGAGCTTCGCACACTTTTATGATTATACCTTTGCGATTGAT 229
DB 184 AGACACACGAGGTGCGTCACTGACATCCCCACATCGGAACATACTCGGAGCCATTGAC 243

QY 230 GGAGTTGATGGAATCTCTTTACGACAGATGAGAGGTTATATTAGTGGGGCATAGTCTTGG 289
DB 244 GAAGCTCCTGACCTCATTTGCCAAATGATGAAAAGTTGTGCTGTTGGTCAACAGCTTTGG 303

QY 290 TGGTATGAAATTTGGGACTTGCCTATGGAAGTATCCACAAAGATCTATGCTGCTGTTT 349
DB 304 TGGCTTGAATTTAGCCATAGCCATGGAAGTTTCCGAAAAAATCTCTGTGCTGTATT 363

QY 350 CTGGCTGCTTTCATGCTGATTTCTGTTCACAACTCTCTCTTGTGTTTGGAAACAGTATAA 409
DB 364 CTGACTGCTTTTCATGCGCGACACCGAACACTCACCATCTCTGCTTGGACAGTTTGG 423

QY 410 TGAGCGAGCCGACCCGAGAAATTTGTTGGATCTCAGTATTTTACCAATATGGTTCCCGTGA 469
DB 424 AAGCAACATGCTTCAAGAGCATGGATGGCACCGGAATTCGAACCTTATGTTTCAGACAA 483

QY 470 AGAGCCACTGACATCCATGTTTTCGCGCAAGTTTCTTGGCTCACAACTCTACCAGCT 529
DB 484 TTCCGAGCTGA---GTATGTTTTTGTAGCCCTGACTTTCATGAAGTTGGGTTCTTACCAGCT 540

QY 530 ATGCTCTCTCTGAGGATCTTGCATTAGCATCATCTTGGTGAGACCAAGCTCTTTTGTATT 589
DB 541 TTCTCCAGTTGAGGATCTTGAATCGGATTAATCTTTTATGAGGCCAGGATCGTTATTAT 600

QY 590 GGAAGACCTTATCGAAGCCCAAGTATTTCAGATGAACGGTTTGGATCAGTGAAGAGT 649
DB 601 TAACGATTTATCGAAGATGAAAACTTCTCGGATGAAGATATGGTCTGTCTTCTCGAGT 660

QY 650 TTACATTTGTGCACTGAGGATGAAGCATACAGAAAGATTTCCAGCGATGGCAAAATGCA 709
DB 661 TTTTCATAGTGTGTAAGAGAGGACAAAGCAATTCAGAAAGAACCGCAGAGATGGATGATGA 720

QY 710 CAACATTTGGTGTCACTGMAAGCAATAGAGATTAAAGGTGCTGATCATATGGCAATGCTATG 769
DB 721 TAATTTTCCGGTGAAATTTAGTGTAGGATGGAGGAGACAGATCATATGCCAATGTTCTG 780

QY 770 CGAGCCCAAAAACTTTTGGCCCTCTCTCTTTGGAAATTTGCCCATPAATAACAACATGATCTCT 829
DB 781 CAAGCCTCAGCAACTCAGTGATTACTTCTCTGAAAAATCGCGGACAAATTCGTTTAATCAA 840

QY 830 ACATTATG 837
DB 841 TCTTCATG 848
```

## RESULT 3

```
ABZ13515
ID ABZ13515 standard; DNA; 792 BP.
XX
AC ABZ13515;
XX
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1320.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
```

```
XX Arabidopsis thaliana.
OS WO200216655-A2.
PN 28-FEB-2002.
PD 24-AUG-2001; 2001WO-US026685.
PF 24-AUG-2000; 2000US-0227866P.
PR 26-JAN-2001; 2001US-0264647P.
PX 22-JUN-2001; 2001US-0300111P.
XX (SCRI ) SCRIPPS RES INST.
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX Harper JF, Kreps J, Wang X, Zhu T;
PI WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed and
PX producing plants with increased tolerance to these abiotic stresses.
XX Claim 144; SEQ ID NO 1320; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 792 BP; 217 A; 167 C; 196 G; 212 T; 0 U; 0 Other;
```

```
Query Match      27.1%; Score 292.2; DB 6; Length 792;
Best Local Similarity 63.1%; Pred. No. 4.2e-68;
Matches 485; Conservative 0; Mismatches 278; Indels 6; Gaps 2;

QY 53 AAAACACTTTGTTTGTAGTACATCGGTGCATGCCATGAGGTTGGAGTTGGTACAAAGCTAAA 112
DB 21 ACAACATTTTGTACTAGTACATCGTTCTGTCATGCGGTGGTGTGCTGTGTTACAAAGTTAA 80

QY 113 GCCACTGTAGAGCTGCAGGCCATAGGTTTACAGCCCTTGATTTAGCAGCTTCTGGCAC 172
DB 81 GCCGCTGTAGAGCGGTGGGCCACCGCGTAACCTGCTGGACTTAGCTGCGCTCCGGAAT 140

QY 173 TGA---TTTGAGAAAAATAGAGGAGCTTCGCACACTTTTATGATTATATCTTGGCAATGAT 229
DB 141 AGACACAAACGAGGTGCGATCACTGACATCCCCACATCGCAACAAATACTCGGAGCCATTGAC 200

QY 230 GGAGTTGATGGAATCTCTTTCAGCAGATGAGAGGTTATATTAGTGGGGCATAGTCTTGG 289
DB 201 GAAGCTCTGTGACCTCATTTGCCAAATGATGAAGGTTGTGCTGTTGTTGTCACAGCTTGG 260

QY 290 TGGTATGAAATTTGGGACTTGTCTATGGAAGAGTATCCACAAAAGATCTATGCTGCTGTTT 349
DB 261 TGGCTTGAATTTAGCCATAGCCATAGGAAAGTTTCCGAAAAAATCTCTGTCGCTGATT 320

QY 350 CTGGCTGCTTTCATGCTGATTTCTGTTTCAAACTCTCTCTTGTGTTTGGAAACAGTATAA 409
DB 321 CTTGACTGCTTTCATGCGCGACACCGAACTCTCACCATCTCTGCTCTTGGACAAAGTTGG 380

QY 410 TGAGCGGACCGCAGCCGAGAAATTTGGTATGACTCAGTTTACCAATATGGTTCCCGTGA 469
DB 381 AAGCAACATGCTTCAAGAGCATGGATGGGCACCGAAATTCGAACCTTATGGTTTCAGACAA 440

QY 470 AGAGCCACTGACATCCATGTTTGTGGGCCAAAGTTTCTTGGCTCACAAAGCTCTACCAGCT 529
```

```
Db      441  TTCCGACTGA---GTATGTTTTTATAGCCCTGACTCATGAAGTTGGTCTCTACAGCT 497
Qy      530  ATGCTCTCTCGAGGATCTTGCAATTAGCAATCATATGGTGAGACCAAGCTCTTTGTTTAT 589
Db      498  TTCTCCAGTTGAGGATCTTGAACCTGGATTACTTTTAAATGAGCCAGGATCGTTATTTAT 557
Qy      590  GGAAGACCTATCGAAGGCCAAGTATTTTCACAGATGAACGGTTTGGATCAGTGAAGAGAGT 649
Db      558  TAAACGATTTATCGAAGATGAAAAAATCTTCTCGATGAAGGATATGGTCTGTTCCTCGAGT 617
Qy      650  TTACATTTGTGTCACCTGAGCATAAAGGCATACCAAGAAATCCAGCGATGSCAAATTGA 709
Db      618  TTTCCATAGTGTGTAAGAGGACAAAGCAATTCAGAAGAACCCAGAGATGGATGATTGA 677
Qy      710  CAACATTTGTGTCACCTGAAGCAATAGAGATTAAGGTGCTGATCAGATGGCAATGCTATG 769
Db      678  TAAATTTCCGGTGAATTTAGTGTGATGGAGATGGAGGACAGATCATATGCCAATGTTCTG 737
Qy      770  CGAGGCCCAAAAACCTTTGGCGCTCTCTCTTTGGAAAATTGCCCAATAATAC 818
Db      738  CAAGCCTCAGCAACTCAGTGATTACTTCTCCTGAAAAATCGCGGACAAATTC 786

RESULT 4
AAC42344
ID      AAC42344 standard; DNA; 816 BP.
AC      AAC42344;
XX
DT      17-OCT-2000 (first entry)
XX
DE      Arabidopsis thaliana DNA fragment SEQ ID NO: 35196.
KW      Hybridisation assay; Genetic mapping; Gene expression control;
KW      Protein identification; signal transduction pathway; metabolic pathway;
KW      Promoter; termination sequence; 58.
XX
OS      Arabidopsis thaliana.
XX
EP1033405-A2.
XX
PD      06-SEP-2000.
XX
PF      25-FEB-2000; 2000EP-00301439.
XX
PR      25-FEB-1999; 99US-0121825P.
PR      05-MAR-1999; 99US-0123180P.
PR      09-MAR-1999; 99US-0123548P.
PR      23-MAR-1999; 99US-0125788P.
PR      25-MAR-1999; 99US-0126264P.
PR      29-MAR-1999; 99US-0126785P.
PR      01-APR-1999; 99US-0127462P.
PR      06-APR-1999; 99US-0128234P.
PR      08-APR-1999; 99US-0128714P.
PR      16-APR-1999; 99US-0129845P.
PR      19-APR-1999; 99US-0130077P.
PR      21-APR-1999; 99US-0130449P.
PR      23-APR-1999; 99US-0130510P.
PR      28-APR-1999; 99US-0130891P.
PR      30-APR-1999; 99US-0131449P.
PR      30-APR-1999; 99US-0132048P.
PR      30-APR-1999; 99US-0132407P.
PR      04-MAY-1999; 99US-0132484P.
PR      05-MAY-1999; 99US-0132485P.
PR      06-MAY-1999; 99US-0132486P.
PR      06-MAY-1999; 99US-0132487P.
PR      07-MAY-1999; 99US-0132863P.
PR      11-MAY-1999; 99US-0134256P.
PR      14-MAY-1999; 99US-0134218P.
PR      14-MAY-1999; 99US-0134219P.
PR      14-MAY-1999; 99US-0134221P.
PR      14-MAY-1999; 99US-0134370P.
PR      18-MAY-1999; 99US-0134768P.
PR      19-MAY-1999; 99US-0134941P.
PR      20-MAY-1999; 99US-0135124P.
PR      21-MAY-1999; 99US-0135353P.
PR      24-MAY-1999; 99US-0135629P.
PR      25-MAY-1999; 99US-0136021P.
PR      27-MAY-1999; 99US-0136392P.
PR      28-MAY-1999; 99US-0136782P.
PR      01-JUN-1999; 99US-0137222P.
PR      03-JUN-1999; 99US-0137528P.
PR      04-JUN-1999; 99US-0137502P.
PR      07-JUN-1999; 99US-0137724P.
PR      08-JUN-1999; 99US-0138094P.
PR      10-JUN-1999; 99US-0138540P.
PR      10-JUN-1999; 99US-0138847P.
PR      14-JUN-1999; 99US-0139119P.
PR      16-JUN-1999; 99US-0139452P.
PR      16-JUN-1999; 99US-0139453P.
PR      17-JUN-1999; 99US-0139492P.
PR      18-JUN-1999; 99US-0139454P.
PR      18-JUN-1999; 99US-0139455P.
PR      18-JUN-1999; 99US-0139456P.
PR      18-JUN-1999; 99US-0139457P.
PR      18-JUN-1999; 99US-0139458P.
PR      18-JUN-1999; 99US-0139459P.
PR      18-JUN-1999; 99US-0139460P.
PR      18-JUN-1999; 99US-0139461P.
PR      18-JUN-1999; 99US-0139462P.
PR      18-JUN-1999; 99US-0139463P.
PR      18-JUN-1999; 99US-0139750P.
PR      18-JUN-1999; 99US-0139763P.
PR      21-JUN-1999; 99US-0139817P.
PR      22-JUN-1999; 99US-0139899P.
PR      23-JUN-1999; 99US-0140353P.
PR      23-JUN-1999; 99US-0140354P.
PR      24-JUN-1999; 99US-0140695P.
PR      28-JUN-1999; 99US-0140823P.
PR      29-JUN-1999; 99US-0140991P.
PR      30-JUN-1999; 99US-0141287P.
PR      01-JUL-1999; 99US-0141842P.
PR      01-JUL-1999; 99US-0142154P.
PR      02-JUL-1999; 99US-0142055P.
PR      06-JUL-1999; 99US-0142390P.
PR      08-JUL-1999; 99US-0142803P.
PR      09-JUL-1999; 99US-0142920P.
PR      12-JUL-1999; 99US-0142977P.
PR      13-JUL-1999; 99US-0143542P.
PR      14-JUL-1999; 99US-0143624P.
PR      15-JUL-1999; 99US-0144005P.
PR      16-JUL-1999; 99US-0144085P.
PR      16-JUL-1999; 99US-0144086P.
PR      19-JUL-1999; 99US-0144325P.
PR      19-JUL-1999; 99US-0144331P.
PR      19-JUL-1999; 99US-0144332P.
PR      19-JUL-1999; 99US-0144333P.
PR      19-JUL-1999; 99US-0144334P.
PR      19-JUL-1999; 99US-0144335P.
PR      20-JUL-1999; 99US-0144352P.
PR      20-JUL-1999; 99US-0144632P.
PR      20-JUL-1999; 99US-0144884P.
PR      21-JUL-1999; 99US-0144814P.
PR      21-JUL-1999; 99US-0145086P.
PR      22-JUL-1999; 99US-0145085P.
PR      22-JUL-1999; 99US-0145087P.
PR      22-JUL-1999; 99US-0145089P.
PR      22-JUL-1999; 99US-0145192P.
PR      23-JUL-1999; 99US-0145145P.
PR      23-JUL-1999; 99US-0145218P.
PR      23-JUL-1999; 99US-0145224P.
PR      26-JUL-1999; 99US-0145276P.
PR      27-JUL-1999; 99US-0145913P.
PR      27-JUL-1999; 99US-0145918P.
```



|  |              |  |
|--|--------------|--|
| PR   | 27-JUL-1999; | 99US-0145919P.   |
| PR   | 28-JUL-1999; | 99US-0145951P.   |
| PR   | 02-AUG-1999; | 99US-0146386P.   |
| PR   | 02-AUG-1999; | 99US-0146388P.   |
| PR   | 02-AUG-1999; | 99US-0146389P.   |
| PR   | 03-AUG-1999; | 99US-0147038P.   |
| PR   | 04-AUG-1999; | 99US-0147204P.   |
| PR   | 04-AUG-1999; | 99US-0147302P.   |
| PR   | 05-AUG-1999; | 99US-0147192P.   |
| PR   | 05-AUG-1999; | 99US-0147260P.   |
| PR   | 06-AUG-1999; | 99US-0147303P.   |
| PR   | 06-AUG-1999; | 99US-0147416P.   |
| PR   | 09-AUG-1999; | 99US-0147493P.   |
| PR   | 09-AUG-1999; | 99US-0147935P.   |
| PR   | 10-AUG-1999; | 99US-0148171P.   |
| PR   | 11-AUG-1999; | 99US-0148319P.   |
| PR   | 12-AUG-1999; | 99US-0148341P.   |
| PR   | 13-AUG-1999; | 99US-0148565P.   |
| PR   | 13-AUG-1999; | 99US-0148684P.   |
| PR   | 16-AUG-1999; | 99US-0149368P.   |
| PR   | 17-AUG-1999; | 99US-0149175P.   |
| PR   | 18-AUG-1999; | 99US-0149426P.   |
| PR   | 20-AUG-1999; | 99US-0149722P.   |
| PR   | 20-AUG-1999; | 99US-0149723P.   |
| PR   | 20-AUG-1999; | 99US-0149929P.   |
| PR   | 23-AUG-1999; | 99US-0149302P.   |
| PR   | 23-AUG-1999; | 99US-0149930P.   |
| PR   | 25-AUG-1999; | 99US-0150566P.   |
| PR   | 26-AUG-1999; | 99US-0150884P.   |
| PR   | 27-AUG-1999; | 99US-0151065P.   |
| PR   | 27-AUG-1999; | 99US-0151066P.   |
| PR   | 27-AUG-1999; | 99US-0151080P.   |
| PR   | 30-AUG-1999; | 99US-0151130P.   |
| PR   | 31-AUG-1999; | 99US-0151438P.   |
| PR   | 01-SEP-1999; | 99US-0151930P.   |
| PR   | 07-SEP-1999; | 99US-0152363P.   |
| PR   | 10-SEP-1999; | 99US-0153070P.   |
| PR   | 13-SEP-1999; | 99US-0153758P.   |
| PR   | 15-SEP-1999; | 99US-0154018P.   |
| PR   | 16-SEP-1999; | 99US-0154039P.   |
| PR   | 20-SEP-1999; | 99US-0154779P.   |
| PR   | 22-SEP-1999; | 99US-0155139P.   |
| PR   | 23-SEP-1999; | 99US-0155486P.   |
| PR   | 24-SEP-1999; | 99US-0155659P.   |
| PR   | 28-SEP-1999; | 99US-0156458P.   |
| PR   | 29-SEP-1999; | 99US-0156596P.   |
| PR   | 04-OCT-1999; | 99US-0157117P.   |
| PR   | 05-OCT-1999; | 99US-0157753P.   |
| PR   | 06-OCT-1999; | 99US-0157865P.   |
| PR   | 07-OCT-1999; | 99US-0158029P.   |
| PR   | 08-OCT-1999; | 99US-0158232P.   |
| PR   | 12-OCT-1999; | 99US-0158369P.   |
| PR   | 13-OCT-1999; | 99US-0159293P.   |
| PR   | 13-OCT-1999; | 99US-0159294P.   |
| PR   | 13-OCT-1999; | 99US-0159295P.   |
| PR   | 14-OCT-1999; | 99US-0159329P.   |
| PR   | 14-OCT-1999; | 99US-0159330P.   |
| PR   | 14-OCT-1999; | 99US-0159331P.   |
| PR   | 14-OCT-1999; | 99US-0159637P.   |
| PR   | 18-OCT-1999; | 99US-0159638P.   |
| PR   | 21-OCT-1999; | 99US-0160741P.   |
| PR   | 21-OCT-1999; | 99US-0160767P.   |
| PR   | 21-OCT-1999; | 99US-0160768P.   |
| PR   | 21-OCT-1999; | 99US-0160770P.   |
| PR   | 21-OCT-1999; | 99US-0160814P.   |
| PR   | 21-OCT-1999; | 99US-0160815P.   |
| PR   | 22-OCT-1999; | 99US-0160980P.   |
| PR   | 22-OCT-1999; | 99US-0160981P.   |
| PR   | 25-OCT-1999; | 99US-0160989P.   |
| PR   | 25-OCT-1999; | 99US-0161404P.   |
| PR   | 25-OCT-1999; | 99US-0161405P.   |
| PR   | 25-OCT-1999; | 99US-0161406P.   |
| PR   | 26-OCT-1999; | 99US-0161359P.   |
| PR   | 26-OCT-1999; | 99US-0161360P.   |
| PR   | 26-OCT-1999; | 99US-0161361P.   |
| PR   | 28-OCT-1999; | 99US-0161920P.   |
| PR   | 28-OCT-1999; | 99US-0161992P.   |
| PR   | 28-OCT-1999; | 99US-0161993P.   |
| PR   | 29-OCT-1999; | 99US-0162142P.   |
| Query Match 27.1%; Score 292.2; DB 3; Length 816;              |              |  |
| Best Local Similarity 63.1%; Pred. No. 4.2e-68;                |              |  |
| Matches 485; Conservative 0; Mismatches 278; Indels 6; Gaps 2; |              |  |
| Qy   | 53           | AAAACACATTTGTTTGTAGTACATGGTGCATGCGCATGGAGGTTCGAGTTCGTACAAAGCTAAA 112 |
| Db   | 45           | ACAACATTTTGTACTAGTACATGGTTGTCGCCATGGCGGTGTCGTGCTGCTGCTACAAAGTTAA 104 |
| Qy   | 113          | GCCACTGCTAGAAGCTGCGAGGCCATAAGTTACAGCCCTTGATTTAGCAGCTTCGGCAC 172      |
| Db   | 105          | GCCGCTGCTAGAGCGGTGGGCCACCGCGTAACCTGCTGGACTTAGCTGCCCTCCGGAAT 164      |
| Qy   | 173          | TGA---TTTGAGAAAAATAGAGGAGCTTCGCACACTTTTATGATTATTACTTTTCCCATTTGAT 229 |
| Db   | 165          | AGACACACAGAGGTGATCACTGACATCCCACATGCGAACATTAATCTCGAGCCATTGAC 224      |
| Qy   | 230          | GGAGTTGATGGAATCTCTTTTCAGCAGATGAGAAAGTTATATTAGTGGGGCATAGTCTTGG 289    |
| Db   | 225          | GAAGCTCCTGACCTCATTTGCCAAATGATGAAAAGTTGTGCTCGTTGGTTCACAGCTTTGG 284    |
| Qy   | 290          | TGGTATGAATTGGGACTTGTCTATGAAAAAGTATCCACAAAGATCTATGCTGCTGTTTTT 349     |
| Db   | 285          | TGGCTTGAACCTTAGCCATAGCCATGGAAGAGTTTCCCGAAAAAATCTCTGTGCTGTATT 344     |
| Qy   | 350          | CTTGCTGCTTTTCATGCTGATTCCTTTCACAACTCCTCTCTTTTGGAAACAGTATAA 409        |
| Db   | 345          | CTTGACTGCTTTTCATGCGCGACACCGAACACTCACATCCTTCGCTTGGACAAGTTTGG 404      |
| Qy   | 410          | TGAGCGGACGCCAGCGCGAGAAATTGGTTGGATATCTCAGTCTTTTACCATATGTTTCCCTGA 469  |
| Db   | 405          | AAGCAACATGCTCTCAAGAAGCATGATGGGCACCGAATTCGAACCTTATGGTTTCAGACAA 464    |
| Qy   | 470          | AGAGCCACTGACATCCATGTTTTTTTGGGCCAAAAGTTCTTGGCTCACAGCTCTACAGCT 529     |
| Db   | 465          | TTCCGGACTGA---GTATGTTTTTTTAGCCCTGACTTCATGAAAGTTGGGTCTCTACCACT 521    |
| Qy   | 530          | ATGCTCTCTCAGAGATCTTGCAATTAGCATCATATTGTTGAGACCAAGCTCTTGTGTTAT 589     |
| Db   | 522          | TTCTCCAGTTGAGGATCTTGAACCTGGGATTACTTTTAATGAGCCAGGATCGTTATTTAT 581     |
| Qy   | 590          | GGAAGACCTATCGAAGGCCAAGTATTTTCACAGATGAACGGTTTGGATCAGTGAAGAGCT 649     |
| Db   | 582          | TAAACGATTTATCGAAGATGAAAAAATCTTCTCGATGAGGATATGGTCTGTTCTCGAGT 641      |
| Qy   | 650          | TTACATTTGTGCTACTGAGGATAAAGGCATACCAAGAAATTCACGCGATGGCAATTGA 709       |
| Db   | 642          | TTTCATAGTGTGTAAGAGGACAAAGCAATTCAGAGAAACCCAGAGATGGATGATTGA 701        |
| Qy   | 710          | CAACATTTGGTCTCACTGAAGCAATAGAGATTAAGGTCTGATCATCATGGCAATGCTATG 769     |
| Db   | 702          | TAAATTTCCGGTGAATTTAGTATGGAGATGGAGAGACAGATCATATATGCCAATGTTCTG 761     |
| Qy   | 770          | CGAGCCCAAAAATCTTGGCCTCTCTCTTGGAAATTCGCCATAAATAC 818                  |
| Db   | 762          | CAAGCCTCAGCACTCAGTGATTACTTCTGAAAATCGGGACAAATTC 810                   |
| RESULT 5   |              |  |
| ADA26531   |              |  |
| ID   | ADA26531     | standard; DNA; 950 BP.   |
| XX   | ADA26531;    |  |
| AC   | ADA26531;    |  |
| XX   | 20-NOV-2003  | (first entry)  |
| DT   | 20-NOV-2003  | (first entry)  |
| XX   |              |  |





|  |              |  |     |
|--|--------------|--|-----|
| PR   | 21-OCT-1999; | 99US-0160767P.   |     |
| PR   | 21-OCT-1999; | 99US-0160768P.   |     |
| PR   | 21-OCT-1999; | 99US-0160770P.   |     |
| PR   | 21-OCT-1999; | 99US-0160814P.   |     |
| PR   | 21-OCT-1999; | 99US-0160815P.   |     |
| PR   | 22-OCT-1999; | 99US-0160980P.   |     |
| PR   | 22-OCT-1999; | 99US-0160981P.   |     |
| PR   | 22-OCT-1999; | 99US-0160989P.   |     |
| PR   | 25-OCT-1999; | 99US-0161404P.   |     |
| PR   | 25-OCT-1999; | 99US-0161405P.   |     |
| PR   | 25-OCT-1999; | 99US-0161406P.   |     |
| PR   | 26-OCT-1999; | 99US-0161359P.   |     |
| PR   | 26-OCT-1999; | 99US-0161360P.   |     |
| PR   | 26-OCT-1999; | 99US-0161361P.   |     |
| PR   | 28-OCT-1999; | 99US-0161920P.   |     |
| PR   | 28-OCT-1999; | 99US-0161992P.   |     |
| PR   | 28-OCT-1999; | 99US-0161993P.   |     |
| PR   | 29-OCT-1999; | 99US-0162142P.   |     |
| Query Match 25.3%; Score 273; DB 3; Length 870;                |              |  |     |
| Best Local Similarity 60.3%; Pred. No. 6.3e-63;                |              |  |     |
| Matches 488; Conservative 0; Mismatches 315; Indels 6; Gaps 2; |              |  |     |
| QY   | 13           | AAAAGAACTACACAGGCATAAAATCAATGAGGAAGGAAACACATTTGTTTAGTAC        | 72  |
| DB   | 1            | AAAAAGTACGAAAGGAAATATGAGTGAGGAGAGAGGAAGCAACATTTCTGCTAGTAC      | 60  |
| QY   | 73           | ATGGTGCATGCCATGAGGTTGGAGTTGGTACAAGCTAAAGCCATGCTAGAGCTGCAG      | 132 |
| DB   | 61           | ATGGTGGTGGCCACGGCGATGGTGTGTGTAACAGGTTAAGCCTCTTCTCGAGGCTTGG     | 120 |
| QY   | 133          | GCCATAAGGTTACAGCCCTTGATTTTAGCAGCTTCTGGCACTGA---TTTGAGAAAAATAG  | 189 |
| DB   | 121          | GCCATCGTTAAACCGCCTTAGACCTAGCTCTCCGGTATAGACACAACACAGTCAATCA     | 180 |
| QY   | 190          | AGGAGCTTCCGACACTTTATGATATATCTTTGCCATTGATGGATGATGGAATCTCTTT     | 249 |
| DB   | 181          | CTGACATTTCTACATGTGAACAAATATCTGAGCCATTGATGAGCTTAATGATCTATTGC    | 240 |
| QY   | 250          | CAGCAGATGAGAAGGTTATATAGTGGGCGATAGTCTTGGTGGTATGAAATTTGGGACTTG   | 309 |
| DB   | 241          | CGAATCATGAGAAGTTGTACTCGTTGGTTCATAGCTTTGGAGGTTTGAGTTAGCCTTAG    | 300 |
| QY   | 310          | CTATGAAAAGTATCCAAAAAGATCTATGCTGCTGTTTCTTGGCTGCTTTTCATGCCCTG    | 369 |
| DB   | 301          | CCATGGATAAGTTTCCCGATAAAATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT    | 360 |
| QY   | 370          | ATTCCTGTTCACAACT   | 429 |
| DB   | 361          | ACACCAAACTCACCATCGTTGTCGAGGAAAAGTTTGAAGCAGCATGACACCAAGAG       | 420 |
| QY   | 430          | ATTGGTTGGATCTCAGTCTTTTACCATATAGTGTTCCTCGAAGAGCCACTGACATCCATGT  | 489 |
| DB   | 421          | GATGGATGGCTCTGAGCTCGACACATATGGTT--CAGATAATTCGGCTTGTCTGTGT      | 477 |
| QY   | 490          | TTTTTGGCCCAAGTTCTTTGGTCAACAAGCTCTCAACAGTATGCTCTCTCTGAGGATCTTG  | 549 |
| DB   | 478          | TCTTCAGCACCGACTTTCATGAGCACCGTCTCTACCAACTTCTCTCTGAGGAGTCTTG     | 537 |
| QY   | 550          | CATTAGCATCATTTGGTGGAGCAAGCTCTTTGTTTATGGAAGACCTTATCGAAGGCCA     | 609 |
| DB   | 538          | AGCTTGGATTGCTTCTTAAAGAGCGCCTAGTTCATTTGTTTATTAATGAATTCGAAGAG    | 597 |
| QY   | 610          | AGTATTTCAGATGAACGGTTTGGATCAGTGAGAGAGTTTACATTTGTGCTGACTGAGG     | 669 |
| DB   | 598          | AGNACTTTTCTGAGAAGGGTATGGATCTGTTCTCTGAGCTTACATTTGTGCAAGAGG      | 657 |
| QY   | 670          | ATAAAGGCATACAGAAGAAATTCACGATGGCAAAATGACAAATTTGGTGTCTCATGAA     | 729 |
| DB   | 658          | ACAACATTAATCTCGGAAGACCATCAACGATGGATGATCCATAATTAATCCGGCGAATTTAG | 717 |
| QY   | 730          | CAATAGAGATTAAGGTGCTGATCACAATGCGATGCTATGCGAGCCCAAAACTTTGCG      | 789 |

|          |   |  |     |
|----------|---|--|-----|
| Db       | 718   | TGATTGAGATGGAAGAGACGGATCATATGCCAATGTTTTCGCAACCTCAAGTACTAAGTG | 777 |
| QY       | 790   | CCTCTCTCTTGGAAATGCCCCATAAATAC                                | 818 |
| DB       | 778   | ACCATCTATTGGCAATCGCTGACAATTTC                                | 806 |
| RESULT 7 |   |  |     |
| ID       | AAC47519  | standard; DNA; 864 BP.                                       |     |
| XX       | AAC47519;   |  |     |
| XX       | AC  | AAC47519;  |     |
| DT       | 18-OCT-2000   | (first entry)  |     |
| DE       | Arabidopsis thaliana  | DNA fragment SEQ ID NO: 54118.                               |     |
| KW       | Hybridisation assay; genetic mapping; gene expression control;          |  |     |
| KW       | protein identification; signal transduction pathway; metabolic pathway; |  |     |
| KW       | promoter; termination sequence; ss.                                     |  |     |
| OS       | Arabidopsis thaliana.   |  |     |
| XX       | EP1033405-A2.   |  |     |
| XX       | PD  | 06-SEP-2000.   |     |
| XX       | PF  | 25-FEB-2000; 2000EP-00301439.                                |     |
| XX       | PR  | 25-FEB-1999; 99US-0121825P.                                  |     |
| PR       | 05-MAR-1999;  | 99US-0123180P.   |     |
| PR       | 09-MAR-1999;  | 99US-0123548P.   |     |
| PR       | 23-MAR-1999;  | 99US-0125788P.   |     |
| PR       | 25-MAR-1999;  | 99US-0126264P.   |     |
| PR       | 29-MAR-1999;  | 99US-0126785P.   |     |
| PR       | 01-APR-1999;  | 99US-0127462P.   |     |
| PR       | 06-APR-1999;  | 99US-0128234P.   |     |
| PR       | 08-APR-1999;  | 99US-0128714P.   |     |
| PR       | 16-APR-1999;  | 99US-0129845P.   |     |
| PR       | 19-APR-1999;  | 99US-0130077P.   |     |
| PR       | 21-APR-1999;  | 99US-0130449P.   |     |
| PR       | 23-APR-1999;  | 99US-0130510P.   |     |
| PR       | 23-APR-1999;  | 99US-0130891P.   |     |
| PR       | 28-APR-1999;  | 99US-0131449P.   |     |
| PR       | 30-APR-1999;  | 99US-0132048P.   |     |
| PR       | 30-APR-1999;  | 99US-0132407P.   |     |
| PR       | 04-MAY-1999;  | 99US-0132484P.   |     |
| PR       | 05-MAY-1999;  | 99US-0132485P.   |     |
| PR       | 06-MAY-1999;  | 99US-0132486P.   |     |
| PR       | 07-MAY-1999;  | 99US-0132487P.   |     |
| PR       | 11-MAY-1999;  | 99US-0132863P.   |     |
| PR       | 14-MAY-1999;  | 99US-0134218P.   |     |
| PR       | 14-MAY-1999;  | 99US-0134219P.   |     |
| PR       | 14-MAY-1999;  | 99US-0134221P.   |     |
| PR       | 14-MAY-1999;  | 99US-0134370P.   |     |
| PR       | 18-MAY-1999;  | 99US-0134768P.   |     |
| PR       | 19-MAY-1999;  | 99US-0134941P.   |     |
| PR       | 20-MAY-1999;  | 99US-0135124P.   |     |
| PR       | 21-MAY-1999;  | 99US-0135353P.   |     |
| PR       | 24-MAY-1999;  | 99US-0135629P.   |     |
| PR       | 25-MAY-1999;  | 99US-0136021P.   |     |
| PR       | 27-MAY-1999;  | 99US-0136392P.   |     |
| PR       | 28-MAY-1999;  | 99US-0136782P.   |     |
| PR       | 01-JUN-1999;  | 99US-0137222P.   |     |
| PR       | 03-JUN-1999;  | 99US-0137528P.   |     |
| PR       | 04-JUN-1999;  | 99US-0137502P.   |     |
| PR       | 07-JUN-1999;  | 99US-0137724P.   |     |
| PR       | 08-JUN-1999;  | 99US-0138094P.   |     |
| PR       | 10-JUN-1999;  | 99US-0138540P.   |     |
| PR       | 10-JUN-1999;  | 99US-0138847P.   |     |
| PR       | 14-JUN-1999;  | 99US-0139119P.   |     |
| PR       | 16-JUN-1999;  | 99US-0139452P.   |     |

PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 03-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.

PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 23-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154799P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0158458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 22-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 25.3%; Score 272.6; DB 3; Length 864;

Best Local Similarity 60.8%; Pred. No. 8.1e-63;

Matches 481; Conservative 0; Mismatches 304; Indels 6; Gaps 2;

Qy 31 ATAAATTCAATGAGGAGGAACACTTTGTTTATTAGTACATGTCATGCCATGGAG 90

Db 17 ATATGAGTGAGGAGGAGGAACACTTTGTTTATTAGTACATGTCATGCCATGGAG 76

Qy 91 GTTGGAGTTGGTACAAAGCTAAAGCCACTGTAGAGCTGCAGGCCATAAGGTTTACAGCCC 150

[illegible]

RESULT 8  
ADA26532

ADA26532  
ID ADA26532 standard: DNA: 1152 BP.

|    |     |
|----|-----|
| AD | ADA |
| XX |     |
| AC |     |

AC  
XX  
DT

XX DE Arabidopsis thaliana ACSH2L3 gene homologous to tobacco SABP2 gene.  
XX ZV NOV 2003 (20030905)

XX salicylic acid-binding protein; SABP2; disease resistance; plant;  
KW modulator; tobacco; ds: gene.

OS *Arabidopsis thaliana*:  
XX  
XX  
XX *modulator*, tobacco;

00  
XX  
PN WO2003016551

XX PD 27-FEB-2003.

| Qy   | 483   | 580 |
|--|---|-----|
| TCCATGTTTTTTTGGCCCCAAGTCTCTTGGCTTCACAGCTCTTACCGACTATGCTCTCTCTGAG | TTTTTGGATTTTTTTGGACAGAAATTCATGGCCAGAACTTGATCAGTTGTCTCCAGTCCAA |     |

QY 543 GATCTTGCATTAGCATCATCAITGGTGAGACCAAGCTCTTTGTTTATGGAAGACCTATCG 602  
|||||  
Db 640 GATCTTGAATTCGCGAAATTTGGTGAGGCGCAACCCATTGATTAAAGAAAGATATGGCA 699  
QY 603 AAGGCCAAGTATTTCAAGATGAACCGTTTGGATCAGTGAAGAGAGTTTACATTTGGTGC 662  
|||||  
Db 700 GAGAGAAGAGCTTTCAGTGAGGAAGGATACGGATCCGTTACACGTATATTATTATTGATGC 759  
QY 663 ACTGAGGATAAAGGCATACAGAGAAATTCAGCGATGCGCAATTCACACATTTGGTGTG 722  
|||||  
Db 760 GGAAGGATCTTGTCACCGGAAGATTACAGCGATCGATGATGACGAACCTTCCGCCCA 819  
QY 723 ACTGAAGCATATAGAGATTAAGGTGCTGATCAGATGCGCAATGCTATGCGAGCCGCCAAAAA 782  
|||||  
Db 820 AAGAAGTAAATGAGATCAAGACGCGAGATCATATGCCAATGTTCTCCAGGCTCAACAA 879  
QY 783 CTTTGGCCCTCTCTTTGAAATTTGCCCAATAATACAACCTGA 824  
|||||  
Db 880 CTATGTGCTCTCTCTTGGAGATTGCAAAATAATATGCGCTAA 921

RESULT 9

ABZ13746  
ID ABZ13746 standard; DNA; 792 BP.

XX AC ABZ13746;  
XX

DT 21-JAN-2003 (first entry)

DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1551.

KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX Arabidopsis thaliana.

XX WO200216655-A2.

XX 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US026685.

XX 24-AUG-2000; 2000US-0227866P.

PR 26-JAN-2001; 2001US-0264647P.

PR 22-JUN-2001; 2001US-0300111P.

XX (SCRI ) SCRIPPS RES INST.

PA (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX Harper JF, Kreps J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

PT Identifying a stress condition to which a plant cell has been exposed and  
PT producing plants with increased tolerance to these abiotic stresses.

XX Claim 144; SEQ ID NO 1551; 577pp + Sequence Listing; English.

CC The invention relates to identifying a stress condition to which a plant  
CC cell has been exposed, comprising: (a) contacting nucleic acid  
CC representative of expressed polynucleotides in the plant cell with an  
CC array or probes representative of the plant cell genome; and (b)  
CC detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present sequence is that  
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
CC in methods of the invention. Note: The sequence data for this patent is  
CC not represented in the printed specification but is based on sequence  
CC information supplied to Derwent by the European Patent Office

XX Sequence 792 BP; 210 A; 169 C; 187 G; 226 T; 0 U; 0 Other;

SQ

Query Match 25.1%; Score 271; DB 6; Length 792;  
Best Local Similarity 61.4%; Pred. No. 2.1e-62;  
Matches 471; Conservative 0; Mismatches 290; Indels 6; Gaps 2;  
QY 55 AACACTTTGTTTGTAGTACATGCGCATGCGCATGGAGTTGGAGTTGGTACAGAGCTAAAGC 114  
Db 23 AACACTTGTGTGTAGTACATGCGTGGTCCACGCGGCATGGTGTGTGTAACAAGGTTAAGC 82  
QY 115 CACTGCTAGAAGCTCGAGGCCATAGGTTTACAGCCCTTGATTAGCAGCTTCTGGCACATG 174  
Db 83 CTCCTTCTCGAGCTTTGGGCCATCGTTAAACCGCTTAGACCTAGCTGCTCCGGTATAG 142  
QY 175 A---TTTGAGAAAAATAGAGGAGCTTCGCACACTTTTATGATTATATCTTTGGCAATGATGG 231  
Db 143 ACACAACAGGTCATCACTGACATTTCTACATGTGAACAAATATTTCTGAGCCATTGATGC 202  
QY 232 AGTTGATGAATCTCTTTTACGACGATGAGAGGTTATATTAGTGGGGCATGATCTTGGTG 291  
Db 203 AGCTAATGACTTTCATTTGCCGAATGATGAGAAGTTGTACTCGTTGGTCATAGCTTTGGAG 262  
QY 292 GTATCAATTTGGGACTTCTGATGGAAGATATCCAAAAAGATCTCAAAAGATCTATGCTGCTTTTCT 351  
Db 263 GTTTCAGTTTACCTTTAGCCATGGATAAGTTTCCGATATAAATCTCTGCTCTGCTTTCG 322  
QY 352 TGGCTGCTTTTCAATGCTGATTTCTGTTTCAAACTCTCTCTTTGTTTGGAAAGATATAATG 411  
Db 323 TGAAGTCAATTCATGCGCGACACCAAACTCAACCTGTTTCGTCGAGGAAGTTTGCAG 382  
QY 412 AGCGAGCGCAGCGAGAAATTTGGTTGGATCTCAGTCTTTTACCATATGTTTCCCTCGAAG 471  
Db 383 GCAGCATGACACACAGAGGATGGATGGGCTCTGAGCTCGAGACATATGGTT---CAGATA 439  
QY 472 AGCCACTGACATCCATGTTTTTGGCCCAAGTTCTTGGCTCACAAGCTCTTACCAGCTAT 531  
Db 440 ATTCCGGCTTGTCTGTGTTCTTCAGCAGCGACTTCATGAAGCAGCTCTCTACCAACTTT 499  
QY 532 GCTCTCTGAGGATCTTGCATTTAGCATCATCATTTGGTGAGACCAAGCTCTTTGTTTATGG 591  
Db 500 CTCCTGTGGAGATCTTGAGCTTGGATTGCTTCTAAAGAGGCTTAGTTTCATTGTTTATTA 559  
QY 592 AAGACCTATCGAAGGCCAAAGTATTTTACAGATGAACGGTTTGGATCAGTGAAGAGATTT 651  
Db 560 ATGAATTTATCGAAGATGGAGAACTTTTCTGAGAAAGGGTATGGATCTGTTCTCGAGCTT 619  
QY 652 ACATTGTGTCACCTGAGCATAAAGCATACAGAGAAATTCAGCGATGGCAAAATGACA 711  
Db 620 ACATTTGTGTGCAAGAGGACACATTTATCTCGAAGAGCCATCAACGATGGATGCCATA 679  
QY 712 ACATTTGGTGTCACTGAAGCAATAGAGATTAAGGTGCTGATCACAATGGCAATGCTATGG 771  
Db 680 ATTATCCGGCGAATTTAGTATTGAGATGGAAGAGACTGATCATATGCCAATGTTTGA 739  
QY 772 AGCCCCAAAACTTTGGCCCTCTCTTTGGAAATTTGCCCATATAATAC 818  
Db 740 AACCTCAACTACTAAGTAGCCATCTATTGGCAATCGCTGACAAATTC 786

RESULT 10

AAC43133  
ID AAC43133 standard; DNA; 792 BP.

XX AC AAC43133;  
XX

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 38151.

XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.

XX Arabidopsis thaliana.

OS

|    |                               |    |              |                |
|----|-------------------------------|----|--------------|----------------|
| PN | EP1033405-A2.                 | PR | 01-JUL-1999; | 99US-0142154P. |
| XX |                               | PR | 02-JUL-1999; | 99US-0142055P. |
| PD |                               | PR | 06-JUL-1999; | 99US-0142390P. |
| XX | 06-SEP-2000.                  | PR | 08-JUL-1999; | 99US-0142803P. |
| PF |                               | PR | 09-JUL-1999; | 99US-0142920P. |
| XX | 25-FEB-2000; 2000EP-00301439. | PR | 12-JUL-1999; | 99US-0142977P. |
| PF |                               | PR | 13-JUL-1999; | 99US-0143542P. |
| PR | 25-FEB-1999;                  | PR | 14-JUL-1999; | 99US-0143624P. |
| PR | 05-MAR-1999;                  | PR | 15-JUL-1999; | 99US-0144005P. |
| PR | 09-MAR-1999;                  | PR | 16-JUL-1999; | 99US-0144085P. |
| PR | 23-MAR-1999;                  | PR | 16-JUL-1999; | 99US-0144086P. |
| PR | 25-MAR-1999;                  | PR | 19-JUL-1999; | 99US-0144325P. |
| PR | 29-MAR-1999;                  | PR | 19-JUL-1999; | 99US-0144331P. |
| PR | 01-APR-1999;                  | PR | 19-JUL-1999; | 99US-0144332P. |
| PR | 06-APR-1999;                  | PR | 19-JUL-1999; | 99US-0144333P. |
| PR | 08-APR-1999;                  | PR | 19-JUL-1999; | 99US-0144334P. |
| PR | 16-APR-1999;                  | PR | 19-JUL-1999; | 99US-0144335P. |
| PR | 19-APR-1999;                  | PR | 20-JUL-1999; | 99US-0144352P. |
| PR | 21-APR-1999;                  | PR | 20-JUL-1999; | 99US-0144632P. |
| PR | 23-APR-1999;                  | PR | 20-JUL-1999; | 99US-0144884P. |
| PR | 23-APR-1999;                  | PR | 21-JUL-1999; | 99US-0144814P. |
| PR | 28-APR-1999;                  | PR | 21-JUL-1999; | 99US-0145086P. |
| PR | 30-APR-1999;                  | PR | 21-JUL-1999; | 99US-0145088P. |
| PR | 30-APR-1999;                  | PR | 22-JUL-1999; | 99US-0145085P. |
| PR | 04-MAY-1999;                  | PR | 22-JUL-1999; | 99US-0145087P. |
| PR | 05-MAY-1999;                  | PR | 22-JUL-1999; | 99US-0145089P. |
| PR | 06-MAY-1999;                  | PR | 22-JUL-1999; | 99US-0145192P. |
| PR | 06-MAY-1999;                  | PR | 23-JUL-1999; | 99US-0145145P. |
| PR | 07-MAY-1999;                  | PR | 23-JUL-1999; | 99US-0145218P. |
| PR | 11-MAY-1999;                  | PR | 23-JUL-1999; | 99US-0145224P. |
| PR | 14-MAY-1999;                  | PR | 26-JUL-1999; | 99US-0145276P. |
| PR | 14-MAY-1999;                  | PR | 27-JUL-1999; | 99US-0145913P. |
| PR | 14-MAY-1999;                  | PR | 27-JUL-1999; | 99US-0145918P. |
| PR | 14-MAY-1999;                  | PR | 27-JUL-1999; | 99US-0145919P. |
| PR | 18-MAY-1999;                  | PR | 28-JUL-1999; | 99US-0145951P. |
| PR | 19-MAY-1999;                  | PR | 02-AUG-1999; | 99US-0146386P. |
| PR | 20-MAY-1999;                  | PR | 02-AUG-1999; | 99US-0146388P. |
| PR | 21-MAY-1999;                  | PR | 02-AUG-1999; | 99US-0146389P. |
| PR | 24-MAY-1999;                  | PR | 03-AUG-1999; | 99US-0147038P. |
| PR | 25-MAY-1999;                  | PR | 03-AUG-1999; | 99US-0147204P. |
| PR | 27-MAY-1999;                  | PR | 04-AUG-1999; | 99US-0147302P. |
| PR | 28-MAY-1999;                  | PR | 05-AUG-1999; | 99US-0147192P. |
| PR | 01-JUN-1999;                  | PR | 05-AUG-1999; | 99US-0147260P. |
| PR | 03-JUN-1999;                  | PR | 06-AUG-1999; | 99US-0147303P. |
| PR | 04-JUN-1999;                  | PR | 06-AUG-1999; | 99US-0147416P. |
| PR | 07-JUN-1999;                  | PR | 09-AUG-1999; | 99US-0147933P. |
| PR | 08-JUN-1999;                  | PR | 09-AUG-1999; | 99US-0147935P. |
| PR | 10-JUN-1999;                  | PR | 10-AUG-1999; | 99US-0148171P. |
| PR | 10-JUN-1999;                  | PR | 11-AUG-1999; | 99US-0148319P. |
| PR | 14-JUN-1999;                  | PR | 12-AUG-1999; | 99US-0148341P. |
| PR | 16-JUN-1999;                  | PR | 13-AUG-1999; | 99US-0148565P. |
| PR | 16-JUN-1999;                  | PR | 13-AUG-1999; | 99US-0148684P. |
| PR | 17-JUN-1999;                  | PR | 16-AUG-1999; | 99US-0149368P. |
| PR | 18-JUN-1999;                  | PR | 17-AUG-1999; | 99US-0149175P. |
| PR | 18-JUN-1999;                  | PR | 18-AUG-1999; | 99US-0149426P. |
| PR | 18-JUN-1999;                  | PR | 20-AUG-1999; | 99US-0149722P. |
| PR | 18-JUN-1999;                  | PR | 20-AUG-1999; | 99US-0149723P. |
| PR | 18-JUN-1999;                  | PR | 20-AUG-1999; | 99US-0149929P. |
| PR | 18-JUN-1999;                  | PR | 23-AUG-1999; | 99US-0149902P. |
| PR | 18-JUN-1999;                  | PR | 23-AUG-1999; | 99US-0149930P. |
| PR | 18-JUN-1999;                  | PR | 25-AUG-1999; | 99US-0150566P. |
| PR | 18-JUN-1999;                  | PR | 26-AUG-1999; | 99US-0150884P. |
| PR | 18-JUN-1999;                  | PR | 27-AUG-1999; | 99US-0151065P. |
| PR | 18-JUN-1999;                  | PR | 27-AUG-1999; | 99US-0151066P. |
| PR | 18-JUN-1999;                  | PR | 27-AUG-1999; | 99US-0151080P. |
| PR | 21-JUN-1999;                  | PR | 30-AUG-1999; | 99US-0151303P. |
| PR | 22-JUN-1999;                  | PR | 31-AUG-1999; | 99US-0151438P. |
| PR | 23-JUN-1999;                  | PR | 01-SEP-1999; | 99US-0151930P. |
| PR | 23-JUN-1999;                  | PR | 07-SEP-1999; | 99US-0152363P. |
| PR | 24-JUN-1999;                  | PR | 10-SEP-1999; | 99US-0153070P. |
| PR | 28-JUN-1999;                  | PR | 13-SEP-1999; | 99US-0153758P. |
| PR | 29-JUN-1999;                  | PR | 15-SEP-1999; | 99US-0154018P. |
| PR | 30-JUN-1999;                  | PR | 16-SEP-1999; | 99US-0154039P. |
| PR | 01-JUL-1999;                  | PR |              |                |



|    |              |                |
|----|--------------|----------------|
| PR | 20-SEP-1999; | 99US-0154779P. |
| PR | 22-SEP-1999; | 99US-0155139P. |
| PR | 23-SEP-1999; | 99US-0155486P. |
| PR | 24-SEP-1999; | 99US-0155659P. |
| PR | 28-SEP-1999; | 99US-0156458P. |
| PR | 29-SEP-1999; | 99US-0156596P. |
| PR | 04-OCT-1999; | 99US-0157117P. |
| PR | 05-OCT-1999; | 99US-0157753P. |
| PR | 06-OCT-1999; | 99US-0157865P. |
| PR | 07-OCT-1999; | 99US-0158029P. |
| PR | 08-OCT-1999; | 99US-0158232P. |
| PR | 12-OCT-1999; | 99US-0158369P. |
| PR | 13-OCT-1999; | 99US-0159293P. |
| PR | 13-OCT-1999; | 99US-0159294P. |
| PR | 13-OCT-1999; | 99US-0159295P. |
| PR | 14-OCT-1999; | 99US-0159329P. |
| PR | 14-OCT-1999; | 99US-0159330P. |
| PR | 14-OCT-1999; | 99US-0159331P. |
| PR | 14-OCT-1999; | 99US-0159637P. |
| PR | 18-OCT-1999; | 99US-0159584P. |
| PR | 21-OCT-1999; | 99US-0160741P. |
| PR | 21-OCT-1999; | 99US-0160767P. |
| PR | 21-OCT-1999; | 99US-0160768P. |
| PR | 21-OCT-1999; | 99US-0160770P. |
| PR | 21-OCT-1999; | 99US-0160814P. |
| PR | 21-OCT-1999; | 99US-0160815P. |
| PR | 22-OCT-1999; | 99US-0160980P. |
| PR | 22-OCT-1999; | 99US-0160981P. |
| PR | 22-OCT-1999; | 99US-0160989P. |
| PR | 25-OCT-1999; | 99US-0161404P. |
| PR | 25-OCT-1999; | 99US-0161405P. |
| PR | 25-OCT-1999; | 99US-0161406P. |
| PR | 26-OCT-1999; | 99US-0161559P. |
| PR | 26-OCT-1999; | 99US-0161360P. |
| PR | 28-OCT-1999; | 99US-0161361P. |
| PR | 28-OCT-1999; | 99US-0161920P. |
| PR | 28-OCT-1999; | 99US-0161992P. |
| PR | 28-OCT-1999; | 99US-0161993P. |
| PR | 29-OCT-1999; | 99US-0162142P. |

Query Match

Best Local Similarity 24.9%; Score 269; DB 3; Length 792;

Mismatches 472; Conservative 0; Mismatches 295; Indels 6; Gaps 2;

|    |     |   |     |
|----|-----|---|-----|
| QY | 55  | AACACITTTGTTTGTAGTACATGGTGCATGGCATGGAGGTTGGAGTTGGTATGGTACAGCTTAAAGC | 114 |
| DB | 23  | AACACGTCGTTCTTAGTACATGGTGTGCTTGGCATGGCGCCTGGTGTGGTACAGGTTAAGC       | 82  |
| QY | 115 | CACGTCTAGAGCTGCAGGCCATAAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCACATG        | 174 |
| DB | 83  | CGCAGCTCGAGGCTTCTGGCCACCGGTAAACCCGCTAGATCTAGCTGCTCCGGTATAG          | 142 |
| QY | 175 | ATTGTG---AGAAAAATAGAGGAGCTTCGCACACTTTATGATTATATCTTTGCCATTGATCG      | 231 |
| DB | 143 | ACATGACCGGTTCAATCAGATATATCCATGCGCAACATCTCAGAGCCATTGATGC             | 202 |
| QY | 232 | AGTTGATGGAATCTCTTTAGCAGATGAGAAGTTATATAGTGGGCAATAGTCTTGGTG           | 291 |
| DB | 203 | AGCTAATGACCTCACTACACAGATGATGAGAAGGTTGTGCTTGTGTCTATAGCTTAGGAG        | 262 |
| QY | 292 | GTATGAAATTTGGCACTTGTATGGAAGTATCCAAAGATCTATGCTGCTGTTTCT              | 351 |
| DB | 263 | GTTTGAAGTTAGCTATGGCAATGGATATGTTTCCGACCAAAATCTCTGTTCTGTCTTTG         | 322 |
| QY | 352 | TGCTGCTTTCATGCTGCTGATTCGTTTCAAACTCTCTCTTTGTTTGGAAACAGTATAATG        | 411 |
| DB | 323 | TGACTGCTATGATGCGACAGACCAAACTCTCACCATCTCTGATGGGATAGCTAAGAA           | 382 |
| QY | 412 | AGCGGAGCGGCGGAGAAATTTGGTGGATCTCAGTTTATTTTACCATATGTTTCCCTGTAAG       | 471 |
| DB | 383 | AGAAAACTTCACGAGGAGGAATGTTAGACACCGTGTTT---ACGAGCGGAAACCTGATT         | 439 |

|    |     |  |     |
|----|-----|--|-----|
| QY | 472 | AGCCACTGACATCCATGTTTTTTTGGCCCAAGTTCTTGGCTCACAAGCTCTTACCAGCTAT  | 531 |
| DB | 440 | TTCTTAGCGAGTTTTTGGATTTTTTGGACCAAGATTTCATGCGCAAGACTTTGTATCACTGT | 499 |
| QY | 532 | GCTCTCTGAGGATCTTGCATTTAGCATCATCATTTGGTGAGACCAAGCTCTTTGTTTATGG  | 591 |
| DB | 500 | CTCCAGTCCAGATCTTGAATTTGGCGAAATGTTGGTGAGGCGCAACCCATTTGATTAAGA   | 559 |
| QY | 592 | AAGACCTATCGAAGGCCAAGTATTTTTCACAGATGAACGGTTTGGATCAGTGAAGAGAGTTT | 651 |
| DB | 560 | AAGATATGCGCAGAGAGAAGAGCTTTCAGTGAGAAGATACGATCCGTTACACGTATAT     | 619 |
| QY | 652 | ACATTGTGTGCACTCAGGATAAAGGCATACCAAGAGAAATTCAGCGATGCGCAATTTGACA  | 711 |
| DB | 620 | TTATTGTATGCGGAAAGGATCTTGTCTACCCGGAAGATTACCAGCGATCGATGATCAGCA   | 679 |
| QY | 712 | ACATTGGTGTCACTCAGCAATAGAGATTAAAGGTGCTGATCAGATGGCAATGCTATGG     | 771 |
| DB | 680 | ACTTTCCCGCAAAAGAGTAATGGAGATCAAGAGCGCAGATCATATGCCAATGTTCTCCA    | 739 |
| QY | 772 | AGCCCCCAAAAGCTTTGGCGCTCTCTCTTGGAAATTTGCCCATAAATACAACTGA        | 824 |
| DB | 740 | AGCCTCAACAACTATGTGCTCTTCTCTTGGAGATTGCAAAATAATATGCTCTAA         | 792 |

RESULT 11

ABZ12412 standard; DNA; 792 BP.

AC ABZ12412;

DT 21-JAN-2003 (first entry)

XX Arabidopsis thaliana stress regulated gene SEQ ID NO 217.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX Arabidopsis thaliana.

XX Arabidopsis thaliana.

XX WO200216655-A2.

XX 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US026685.

XX 24-AUG-2000; 2000US-0227866P.

PR 26-JAN-2001; 2001US-0264647P.

PR 22-JUN-2001; 2001US-0300111P.

XX (SCRI ) SCRIPPS RES INST.

PA (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX Harper JF, Kreps J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

XX Identifying a stress condition to which a plant cell has been exposed and

XX producing plants with increased tolerance to these abiotic stresses.

XX Claim 144; SEQ ID NO 217; 577pp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant

XX cell has been exposed, comprising: (a) contacting nucleic acid

XX representative of expressed polynucleotides in the plant cell with an

XX array or probes representative of the plant cell genome; and (b)

XX detecting a profile of expressed polynucleotides in the plant cell

XX characteristic of a stress response. The method is useful in the

XX production of transgenic plants, cells and seeds and in producing plants

XX with increased tolerance to abiotic stress. The present sequence is that

XX of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used

XX in methods of the invention. Note: The sequence data for this patent is

XX not represented in the printed specification but is based on sequence

XX information supplied to Derwent by the European Patent Office

```
XX SQ Sequence 792 BP; 230 A; 167 C; 192 G; 203 T; 0 U; 0 Other;
Query Match 24.9%; Score 269; DB 6; Length 792;
Best Local Similarity 61.1%; Pred. No. 7.2e-62;
Matches 472; Conservative 0; Mismatches 295; Indels 6; Gaps 2;
QY 55 AACACTTTGTTTGTAGTACATGTCATGCCATGCCATGGAGTTGGAGTTGGTACAGCTAAAGC 114
Db 23 AACACGTCGTTCTAGTACATGTTGCTTGCATGGCGCCTGCTGGTACAGGTTAAGC 82
QY 115 CACTGCTAGAACGTCGAGGCCATAAGTTTACAGCCCTTGATTATAGCAGCTTCTGGCACTG 174
Db 83 CGCAGCTCGAGCTTCTGGCCACCCGCTAACCGCGTAGACTAGCTGCTCCGATAG 142
QY 175 ATTTG---AGAAAAATAGAGGAGCTTCGCACACTTTATGATTATATCTTTGCCATTGATGG 231
Db 143 ACATGACCAGGTCATACAGATATATCCACATGCAACAATACTCAGAGCCATTGATGC 202
QY 232 AGTTGATGGAATCTCTTTACAGAGATGAGAAGGTTATATTAGTGGGCATAGTCTTGGTG 291
Db 203 AGCTAATGACCTCACTTACCAGATGATGAGAAGGTTGTGCTTGTGTCATAGCTTAGGAG 262
QY 292 GTATGAATTTGGAGCTTGTCTATGGAAGATATCCACAAAGATCTATGCTGCTGTTTCT 351
Db 263 GTTTGAGTTTACGATGGCCATGGATATGTTTCCGACCAAAATCTGTTCTGCTTTG 322
QY 352 TGGCTGCTTTTCATGCTGATCTGTTTCAACAACCTCTCTTTGTTTGGAAACAGTATAATG 411
Db 323 TGACTGCTATGATGCCAGACACCAACACTCAACCTCTCTGTTGGGATAGCTTAAGAA 382
QY 412 AGGGAGCCGACCGAGATTTGGTTGGATATCTAGCTTTTACCATATGTTTCCCTCGAAG 471
Db 383 AAGAAACTTTCAGGAGGAAGTGGTTAGACACCGTGT---ACGAGCGAGAAACCTGATT 439
QY 472 AGCCTGATACATCCATGTTTTTTGGCCCAAGTCTTGGCTCACAAGCTCTACCAAGCTAT 531
Db 440 TTCCTAGCGAGTTTGGATTTTGGACCAAGATTCATGGCCAAAGACTTGTATCAGTTGT 499
QY 532 GCTCTCCTGAGGATCTTGCAATGATGATCATCATTTGGTGAGACCAAGCTCTTTGTTTATGG 591
Db 500 CTCCTAGTCCAGATCTTGAATTTGGCAAAATGTTGGTAGGGCAAAACCCATTGATTAAGA 559
QY 592 AAGACCTATCGAAGCCAGTATTTTACAGATGAACGGTTTGGATGAGTGAAGAGATTT 651
Db 560 AAGATATGCGAGAGAAAGAGCTTCAGTGAGGAAGGATACGGATCCGTTACACGTATAT 619
QY 652 ACATTGTTGTCATGAGGATAAAGGCATACCAAGAAATTCAGCGATGGCAAAATTGACA 711
Db 620 TTATTGTATGCGAAAGGATCTTGTGTACCCGAGAGATTACAGCGATCGATGATCAGCA 679
QY 712 ACATTGGTCTCACTGAAGCAATAGAGATTTAAAGGTGCTGATCACAATGGCAATGCTATGG 771
Db 680 ACTTCCCCCAAAAGAAAGTAAATGAGATCAAGAGCGAGATCATATGCAATGTTCTCCA 739
QY 772 AGCCCCAAAACCTTTGGCGCTCTCTCTTGGAAATTTGCCCATAAATACAACTGA 824
Db 740 AGCCTCAACAACACTATGTGCTCTTCTCTGGAGATTGCAAAATAATATGCTTAA 792
```

RESULT 12  
ADN74620  
ID ADN74620 standard; cDNA; 792 BP.  
XX  
AC ADN74620;  
XX  
DT 15-JUL-2004 (first entry)  
XX  
DE Thale cress cDNA repressed in E2Fa/Dpa expressing plants SeqID 2515.  
XX  
KW gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;  
KW growth regulator; animal feed product; thale cress;  
KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.

```
XX OS Arabidopsis thaliana.
XX FN WO2004035798-A2.
XX PD 29-APR-2004.
XX PF 20-OCT-2003; 2003WO-EP011658.
XX PR 18-OCT-2002; 2002EP-00079408.
XX PA (CROP-) CROPDESIGN NV.
XX PI Inze D, De Veylder L, Vlieghe K;
XX WPI; 2004-348466/32.
XX P-PSDB; ADN74621.
XX Altering plant characteristics, useful for producing plants for enzyme or
XX pharmaceutical production comprises modifying in a plant, expression of
XX one or more nucleic acids and/or modifying level or activity of one or
XX more proteins.
XX Claim 1; SEQ ID NO 2515; 134pp: English.
XX This invention relates to a novel method for altering one or more plant
XX characteristics. Specifically, it refers to identifying genes that are up
XX - or down-regulated in transgenic plants overexpressing the heterodimeric
XX E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
XX alter plant characteristics accordingly. The present invention describes
XX generating transgenic plants for the production of growth regulators,
XX enzymes, therapeutics, pharmaceuticals and animal feed products, where
XX the altered plant characteristics are selected from increased yield or
XX biomass, enhanced survival capacity, stress tolerance, plant architecture
XX or physiology, altered endoreplication, signal transduction, signal
XX transduction, storage lipid mobilisation and/or altered photosynthesis,
XX each relative to the corresponding wild type plants. Accordingly, these
XX sequences can also be useful as positive or negative selectable markers
XX during transformation of cells or tissues. The identified genes play a
XX role in a variety of biological processes such as DNA replication, cell
XX wall biosynthesis, nitrogen and/ or carbon metabolism or they function as
XX transcription factors. This polynucleotide sequence is thale cress cDNA
XX repressed 1.3 fold or more in plants overexpressing the E2Fa/Dpa
XX transcription factor, given in an exemplification of the invention.
XX SQ Sequence 792 BP; 230 A; 167 C; 192 G; 203 T; 0 U; 0 Other;
Query Match 24.9%; Score 269; DB 12; Length 792;
Best Local Similarity 61.1%; Pred. No. 7.2e-62;
Matches 472; Conservative 0; Mismatches 295; Indels 6; Gaps 2;
QY 55 AACACTTTGTTTGTAGTACATGTCATGCCATGCCATGGAGTTGGAGTTGGTACAGCTAAAGC 114
Db 23 AACACGTCGTTCTAGTACATGTTGCTTGCATGGCGCCTGCTGGTACAGGTTAAGC 82
QY 115 CACTGCTAGAACGTCGAGGCCATAGAGTTTACAGCCCTTGATTATAGCAGCTTCTGGCACTG 174
Db 83 CGCAGCTCGAGCTTCTGGCCACCCGCTAACCGCGTAGACTAGCTGCTCCGATAG 142
QY 175 ATTTG---AGAAAAATAGAGGAGCTTCGCACACTTTATGATTATATCTTTGCCATTGATGG 231
Db 143 ACATGACCAGGTCATACAGATATATCCACATGCAACAATACTCAGAGCCATTGATGC 202
QY 232 AGTTGATGGAATCTCTTTACAGAGATGAGAAGGTTATATTAGTGGGCATAGTCTTGGTG 291
Db 203 AGCTAATGACCTCACTTACCAGATGATGAGAAGGTTGTGCTTGTGTCATAGCTTAGGAG 262
QY 292 GTATGAATTTGGAGCTTGTCTATGGAAGATATCCACAAAGATCTATGCTGCTGTTTCT 351
Db 263 GTTTGAGTTTACGATGGCCATGGATATGTTTCCGACCAAAATCTGTTCTGCTTTG 322
QY 352 TGGCTGCTTTTCATGCTGATCTGTTTCAACAACCTCTCTTTGTTTGGAAACAGTATAATG 411
XX
```

Db 323 TGACTGCTATGATGCCAGACACCAACACATCCACCATCTCTTCGTATGGATAAGCTAAGAA 382  
QY 412 AGCGGACGCCAGCGGAGAAATGGTTGGATACTCAGTTTATACCATATGGTTCCCTCGAAG 471  
Db 383 AAGAACTTCACAGAGGAATGGTTAGACACCGTGT---ACGAGGAAACCTGATTT 439  
QY 472 AGGCATGACATCCATGTTTTTGGGCCAAAGTTCTTGGCTCAAAAGCTCTACCGCAT 531  
Db 440 TTCTTAGCGAGTTTGGATTTTGGACCAAGATTCATGGCCAAAGACTTGTATCAGTTGT 499  
QY 532 GCTCTCTGAGGATCTGCAATAGCATCATCTTGTGAGACCAAGCTCTTTGTTATGG 591  
Db 500 CTCAGTCCAAAGATCTTGAATGGCGAAATGTTGTGAGGGCAAAACCCCATGATTAAGA 559  
QY 592 AAGACCTATCCAAAGGCCAAGTATTTACAGATGAACGGTTTGGATCAGTGAAGAGAGTTT 651  
Db 560 AAGATATGGCAGAGAGNAGAGCTTCAGTGAGAGAGATACGGATCCGTTACAGTATAT 619  
QY 652 ACATTTGTGCTGCTGAGGATAAAGGCATACCAAGAAATTCACGCGATGGCAAAATGACA 711  
Db 620 TTATTTGATGCGGAAAGGATCTTGTGTCACCCGAAAGATTACCAAGCATGATCAGCA 679  
QY 712 ACATTTGTGCTGCTGAGGATAAAGGCATACCAAGAAATTCACGCGATGGCAAAATGACA 771  
Db 680 ACTTTCCCAAAAGAGTAATGGAGATCAAGACGCGAGATCATATGCCAATGTTCTCCA 739  
QY 772 AGCCCCAAACCTTTGGCGCTCTCTCTTGGAAATGCCCAATACCAACTGA 824  
Db 740 AGCCTCAACACTATGTGCTCTCTCTTGGAGATTGCAAAATAAATATGCGCTAA 792

## RESULT 13

ADA26533  
ID ADA26533 standard; DNA; 1109 BP.

XX AC ADA26533;

XX XX 20-NOV-2003 (first entry)

DE Arabidopsis thaliana AtS2L4 gene homologous to tobacco SABP2 gene.

KW salicylic acid-binding protein; SABP2; disease resistance; plant;

KW modulator; tobacco; db; gene.

OS Arabidopsis thaliana.

XX W02003016551-A2.

XX XX 27-FEB-2003.

XX XX 16-AUG-2002; 2002WO-US026312.

XX XX 16-AUG-2001; 2001US-0312863P.

XX (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

XX PA Klessig DF, Kumar D;

XX PI WPI; 2003-278578/27.

XX New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for  
PT producing transgenic plants with increased resistance to disease, or for  
PT screening SABP2 modulators that confer enhanced resistance of plants to  
PT disease.

XX Claim 65; Page; 98pp; English.

XX The invention relates to a novel isolated 1079 bp salicylic acid-binding  
CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid  
CC protein, sequences that specifically hybridize with the nucleic acid, the  
CC complement of the nucleic acid or a natural allelic variant of the  
CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for  
CC conferring increased resistance to disease in plants, or for screening

CC modulators of SABP2, which confer increased or enhanced resistance of  
CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also  
CC useful for producing transgenic plants with increased resistance to  
CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful  
CC for studying the molecular mechanisms responsible for the modulation of  
CC salicylic acid-mediated disease resistance in plants. This sequence  
CC represents the gene encoding the Arabidopsis thaliana SABP2 AtS2L4 which  
CC has homology to the gene (ADA26498) encoding the tobacco SABP2 protein.  
CC (Note: this sequence is not given in the specification but is derived  
CC from the Genbank accession number given by the inventors).

XX SQ Sequence 1109 BP; 365 A; 193 C; 235 G; 316 T; 0 U; 0 Other;

Query Match 23.7%; Score 256; DB 8; Length 1109;

Best Local Similarity 58.6%; Pred. No. 2.6e-58;

Matches 464; Conservative 0; Mismatches 325; Indels 3; Gaps 1;

QY 36 ATTCAAATGAAGAGGAAAAACATTTGTTTATGATGCTGGTGCATGCCATGAGGTTGG 95

Db 1 ATGGAGAAGAACAAACAAGAACGGTTTGTCTCGTCCATGGCTGTGCCACGCGCTTGG 60

QY 96 AGTTGGTCAAGCTAAAGCCACTGTAGAGCTGCGAGGCCATAAGTTACACCCCTTGAT 155

Db 61 TGTGGTACAGGTGAAAAACGATCTGGAGGCTGTAGTCTCTGTGTGACCCGCGTGGAT 120

QY 156 TTAGCAGCTTCTGGCAGCTGATTGTGAGAAAAATAGAGGAGCTTCGCACACTTTATGATTAT 215

Db 121 CTAGCTGCATCCGGTATAAATAGTAGATTGGAAAGAGATTCAGACTTTGAAGGATTAC 180

QY 216 ACTTTGCCATTGATGGAGTTGATGGAATCTCTT---TCAGCAGATGAGAAGGTTATATTA 272

Db 181 TCGAAACCTTTGCTTGGTTACTGAACTCAGTTGGCTGGATGACGATAAGGTGATCTT 240

QY 273 GTGGGCGCATAGTCTTGGTGGTATGAATTTGGAGCTGTCTATCGAAAAAGTATCCAAAAAG 332

Db 241 GTTGGCGCATAGTATGGAGGAATACCTGCTGCTCTCTGCTTCTGACATATATCCCTAGTAAG 300

QY 333 ATCTATGCTGCTGTTTCTTGGCTGCTTTTCATGCGCTGATTCCTGTTCCACAACCTCCTCTT 392

Db 301 ATTGCTACTATTGTTTCTTGACAGCTTTTATGCCCGCACAAAGAACTACTCTGCTTAT 360

QY 393 GTTTGGAAACAGTATAATGAGCGGACGCCAGCGAGAAATGGTTGGATACTCAGTTTATTA 452

Db 361 GTTTACCAAAAGCTAATCAGAGCGTTCACAAGAAGGATGCTTGGACACCGGTTTGA 420

QY 453 CCATATGTTCCCTGGAAGAGCCACTGACATCCATGTTTGTGGCCCAAAAGTTCTTGGCT 512

Db 421 ACGTATGGGAAACATGAATGTCTCTAGAGTTTGTCTTTTGGACCACAAAGTTTCATGGCC 480

QY 513 CACAAGCTCTACAGCTATGCTCTCTGAGGATCTTGCATTAGCATCATTTGGTGAGA 572

Db 481 AAGAATTTGTATCAACTCTCTCCGGTCCAAGATCTTGAATTTGCCGAAATATTTGGTGAGA 540

QY 573 CCAAGCTCTTTGTTTATGGAAGACCTATCGAAGGCCAAAGTATTTCCACAGATGAACGGTTT 632

Db 541 GTAAACCCCATCATTTACAAATAATCTGGCAGGGAACAAGAGCTTTAGTGAGGAAGGTAC 600

QY 633 GGATCAGTGAAGAGAGTTTACATTTGTGTGCACTGAGGATTAAGGCGATACCCAGAAATTC 692

Db 601 GGATCCGTTTACACGTATATATATTTGTATGTGGAGAGGACATGGCGGTATCCCGAGGATTAC 660

QY 693 CAGCGATGCGAAATTTGACAACAATTTGGTGTCTACTGAAGCAATAGAGATTAAGGCTGAT 752

Db 661 CAGTGGTGGATGATCAAGAACTTTCCGCCAAAAGAAAGTAAATGGAGATCAAAATGTCAGAT 720

QY 753 CACATGGCAATGCTATGCGAGCCCCCAAAACCTTTGCGGCTCTCTCTTGGAAATGGCCCAT 812

Db 721 CATATGGCAATGTTCTCAAGCCTCACAACCTATGTGCTCTCTCTGCGAGAGATTGCATGT 780

QY 813 AAATACAACTGA 824

Db 781 AAATATGCTAA 792

```

RESULT 14
ADA26535
ID ADA26535 standard; DNA; 1046 BP.
XX AC
XX ADA26535;
XX
XX 20-NOV-2003 (first entry)
XX
XX Arabidopsis thaliana AtSB2L6 gene homologous to tobacco SABP2 gene.
DE
XX salicylic acid-binding protein; SABP2; disease resistance; plant;
KW modulator; tobacco; ds; gene.
XX Arabidopsis thaliana.
OS
XX WO2003016551-A2.
XX
XX 27-FEB-2003.
XX
XX 16-AUG-2002; 2002WO-US026312.
XX
XX 16-AUG-2001; 2001US-0312863P.
XX
XX (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX
XX Klessig DF, Kumar D;
XX
XX WPI; 2003-278578/27.
XX
XX New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for
XX producing transgenic plants with increased resistance to disease, or for
XX screening SABP2 modulators that confer enhanced resistance of plants to
XX disease.
XX
XX Claim 65; Page; 98pp; English.
XX
XX The invention relates to a novel isolated 1079 bp salicylic acid-binding
XX protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid
XX protein, sequences that specifically hybridize with the nucleic acid, the
XX complement of the nucleic acid or a natural allelic variant of the
XX sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for
XX conferring increased resistance to disease in plants, or for screening
XX modulators of SABP2, which confer increased or enhanced resistance of
XX plants to disease. The SABP2 or SABP2 homologue nucleic acid is also
XX useful for producing transgenic plants with increased resistance to
XX disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful
XX for studying the molecular mechanisms responsible for the modulation of
XX salicylic acid-mediated disease resistance in plants. This sequence
XX represents the gene encoding the Arabidopsis thaliana SABP2 AtSB2L6 which
XX has homology to the gene (ADA26498) encoding the tobacco SABP2 protein.
XX (Note: this sequence is not given in the specification but is derived
XX from the Genbank accession number given by the inventors).
XX
XX Sequence 1046 BP; 330 A; 195 C; 248 G; 273 T; 0 U; 0 Other;
XX
XX
XX Query Match 23.4%; Score 252.6; DB 8; Length 1046;
XX Best Local Similarity 58.2%; Pred. No. 2.1e-57;
XX Matches 463; Conservative 0; Mismatches 329; Indels 3; Gaps 1;
XX
XX 26 AGGCGATTAATCAATGAAGGAGGAAACACATTTGTTAGTACATGGTGCATGCCA 85
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX 12 AAGAAAGGCAATGGAGATAAGAACCCAGAACCGGTTTGTGCTCATCGGGGTGGCCA 71
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX 86 TGGAGGTTCGAGTTGCTCAAGCTAAAGCCACTGTAGAAAGCTGAGGCCATGAAGTTAC 145
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX 72 CGGGGATGAGCTTGGGACAGAGGTGAAGAGCGAGCTGGAGTTGAGGTCACTGTGTGAC 131
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX 146 AGCCCTTGATTTAGCAGCTTCTGGCAGTCAATTTGAGAAAAATAGAGAGCTTGCACACT 205
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX 132 GGCAGTGGATCTTGTGTCATCAGGTATAACATGACCAAGTGAAGAGATTGAGACTCT 191
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX 206 TTATGATTACTTCTGCCATTTGATGGAGTTGATGGATCTCTT--TCAGCAGATGAGAA 262
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 192 GAACGATTACTGCAAAACCAATTGCTTGAGTTTCTGAGCTCGCTTGGCTCAGATGACGGTAA 251
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 263 GGTATATATTAGTGGGCAATAGTCTTGGTGGTATGAATTTGGAGCTTCTGCTATGGAAGTA 322
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 252 GGTGATTTGTTGCTCATAGCATGGGAGGAATATCCGCTGCACCTTGTCTGTGACAGCTT 311
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 323 TCCACAAAAGATCTATGCTGCTGCTTTTCTTGGCTGCTTTTCATGCTGATTCCTGTTCAAA 382
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 312 CGCTTGAAGATTCCCGCTATTGCTTTTGGACAGCTTTTCATGCCCGCACAAATAACCC 371
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 383 CTCTCTCTTTGTTTGGAAACAGTATAATGAGCGGACCGCCAGAGAAATGGTTGGATAC 442
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 372 ACCTGCTTATGTTTACGAAAGAGCTGCTCAGAAGCATTCACAGAGGAATGTTGGACAC 431
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 443 TCAGTTTTTACCATATGTTTCCCTGAAGAGCCACTGACATCCATGTTTTTGGCCCAA 502
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 432 CACGTGTGTGAACCTACGGGAAACCTGATTTTCTCTACGTATACTCTTTTGGGACCAA 491
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 503 GTTCTTGGCTCACAGCTCTACAGCTATGCTCTCCTGAGGATCTTGCAATAGCATATC 562
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 492 GTTTATGGCCAAAGAAATGTTATCAAACTCTCCAGTTCAAGATCTTGAAGTGGTGAAGAC 551
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 563 ATTGCTGAGACCAAGCTCTTTGTTTATGGAAGACCTATCGAAGGCCAAGTATTTACAGA 622
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 552 ATTAGTAGGGGAAACCGTTAGTTACAAACAACTCTGGCAGGGACAAAGAAGCTTTAGTGA 611
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 623 TGAACGGTTTGGATCAGTGAAGAGAGTTTACATTTGTGTGCACTGAGGATAAGGCATACC 682
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 612 GGAAGGGTACCGATCCGTTACACGTATATATTTGTATGTCAGAGAGGATCTTGTGGAAGT 671
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 683 AGAAGAAATTCAGCGATGGCAAAATTTGACAACTGGTGTCTACTGAAGCAATAGAGATTA 742
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 672 CGAAGATTACCAGCGTTGGATGATTAGCACTTTCCACCAAAAGAGTAATGGAGATCAA 731
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 743 AGTGCTGATCAGATGCAATGCTATGCGAGCCCCCAAAATTTTTCGCTCTCTCTTGA 802
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 732 ATGTGCAATCATATGCCAATGTTCTCCAAGCTTCAAGAAGTTTGTGCTCTCTCTCGA 791
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 803 AATTGCCCATTAATA 817
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 792 GATTGCAATAATA 806
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 15
ADA26538
ID ADA26538 standard; DNA; 928 BP.
XX AC
XX ADA26538;
XX
XX 20-NOV-2003 (first entry)
XX
XX Arabidopsis thaliana AtSB2L9 gene homologous to tobacco SABP2 gene.
DE
XX salicylic acid-binding protein; SABP2; disease resistance; plant;
KW modulator; tobacco; ds; gene.
XX Arabidopsis thaliana.
OS
XX WO2003016551-A2.
XX
XX 27-FEB-2003.
XX
XX 16-AUG-2002; 2002WO-US026312.
XX
XX 16-AUG-2001; 2001US-0312863P.
XX
XX (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX
XX Klessig DF, Kumar D;
XX
XX WPI; 2003-278578/27.
XX
XX

```

PT New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for  
PT producing transgenic plants with increased resistance to disease, or for  
PT screening SABP2 modulators that confer enhanced resistance of plants to  
PT disease.  
XX  
XX  
PS Claim 65; Page; 98pp; English.

XX The invention relates to a novel isolated 1079 bp salicylic acid-binding

CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid  
CC protein, sequences that specifically hybridize with the nucleic acid, the  
CC complement of the nucleic acid or a natural allelic variant of the  
CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for  
CC conferring increased resistance to disease in plants, or for screening  
CC modulators of SABP2, which confer increased or enhanced resistance of  
CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also  
CC useful for producing transgenic plants with increased resistance to  
CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful  
CC for studying the molecular mechanisms responsible for the modulation of  
CC salicylic acid-mediated disease resistance in plants. This sequence  
CC represents the gene encoding the Arabidopsis thaliana SABP2 AtSB2L9 which  
CC has homology to the gene (ADA26498) encoding the tobacco SABP2 protein.  
CC (Note: this sequence is not given in the specification but is derived  
CC from the Genbank accession number given by the inventors).

XX SQ Sequence 928 BP; 279 A; 174 C; 214 G; 261 T; 0 U; 0 Other;

Query Match 22.5%; Score 243.2; DB 8; Length 928;  
Best Local Similarity 57.0%; Pred. No. 6.7e-55;  
Matches 455; Conservative 0; Mismatches 348; Indels 3; Gaps 1;

QY 11 AGAAAGAACTAACAGGCGATAAATTCAAATGAAGGAAGGAAAAACATTTGTTTGTAGT 70  
DB 11 AAAAAAAGGCAAACTACAGAAAGAAAGAAAGGAGAGATGAAGCATATGTGCTAGT 70  
QY 71 ACATGTGCATGCCATGGAGTTGGAGTTGGTACAGCTTAAGCCACTGCTAGAGCTGC 130  
DB 71 TCACGGAGGCTGCCACGGTGCCTGTTGGTCAAGGTGAAGCCGATGCTTGAACATTC 130  
QY 131 AGGCCATAAGTTACAGCCCTTCATTTAGCAGCTTCTGGCAGCTGATTCGAGAAAAATAGA 190  
DB 131 CGGCCACCGTGCACGGTTTTTGATCTTAGCGGCGATGGTGAACATGAGCAGAGTAGA 190  
QY 191 GGAGCTTCGCACACTTTATGATATATCTTTGCCAATTGATGGAGTTGATGGAATCTCTTTC 250  
DB 191 AGATATTCAGACTTTGGAGGATTTTCGTAAGCCGTTGCTTGAGTTCTTGAGTCTTTTGG 250  
QY 251 AGCAGATGAGAAGTTATATAGTGGGCGATAGTCTTTGGTGGTATGAAATTTGGGACTTGC 310  
DB 251 CTCGGATGATAAAGTAGTCTCTCGCGCATAGCTCGGTGGAAATACCGGCTGCTCTTGC 310  
QY 311 TATGGAAGATATCCACAAAGATCTATGCTGCTGTTTCTTGGCTGTTTCATGCTGA 370  
DB 311 AGCCGACATGTTTCTAGTAAATATCTGTGTGCTGCTTCTGTTACTTCTTTATGCCGA 370  
QY 371 TTCTGTTCACACTCCCTCTTTGTTTGGAAAGTATAATGAGCGGACGCCAGCCAGAA 430  
DB 371 CACAACGAATCCACCTTCTTACGTGTTTCGAAAAGTTTCTCGAAGCATTACAGAGAAGA 430  
QY 431 TTGGTTGGATACATGTTTACCATATGTTTCCCTGAAAGAGCCACTGACATCCATGTT 490  
DB 431 ACGTATGGACTTCGAGTTAGGGAGCTATGGAACAGATGACCATCCACTAAAGACTGCTTT 490  
QY 491 TTTTGGCCCAAGTCTCTGGCTCACAAGCTCTACAGCTATGCTCTCTGAGGATCTTGC 550  
DB 491 TCTTGGACCTTAACTCTTG---AAGAAATGATATCTTCTCTCTATCGAAGATTATGA 547  
QY 551 ATTAGCATCATATTGGTGAGACCAAGCTCTTTGTTTATGGAAGACTTATCGAAGGCCAA 610  
DB 548 ATTGGCCAAAATGTTGATGAGATCACACGGCTATTACTAGTAATCTGACGGGACTAA 607  
QY 611 GTATTTTCAGATGAACGGTTTGGATCAGTGAAGAGATTTTACATTTGTGTGCACTGAGGA 670  
DB 608 AAGCTTAAACGGCACAGGATATGGATCGATTAGTCGTGTGTATATCGTATGCGGAGAAGA 667

QY 671 TAAAGGCATACCAAGAAGATTCACGCGATGGCAAAATTGACAAACATTTGGTGTCACTGAAGC 730  
DB 668 TAAGGGTATACGTGTAGATTTCCACCGATGGATGATTGAGAACTCTCCGGTTAAAGAAGT 727  
QY 731 AATAGAGATTAAAGGTGCTGATCATATGGCAATGCTATCGGAGCCCAAAAACTTTGCGC 790  
DB 728 GATGGAGATCAAAAGATGCAGATCATATGCTATGTTTCCAAAGCCTCATGAACCTCTGTGA 787  
QY 791 CTCTCTCTTGGAAATTTGCCATAAATAACAACCTGATC 826  
DB 788 TCGTCTTTTAAAGATTGCTGATAAATATATCCCTAAGC 823

Search completed: September 12, 2005, 02:21:38  
Job time : 597.255 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 00:50:21 ; Search time 3440.23 Seconds  
(without alignments)  
11938.557 Million cell updates/sec

Title: US-10-780-002-1

Perfect score: 1079

Sequence: 1 acgcggggaagaagaaaa.....gaaaaaaaaaaaaaaaaa 1079

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsa1:\*  
9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|----|-------------|
| 1          | 575.6 | 53.3  | 922    | 7  | CK269872    |
| 2          | 566.6 | 52.5  | 886    | 7  | CK260266    |
| 3          | 562.1 | 52.1  | 914    | 7  | CK270870    |
| 4          | 562   | 52.1  | 966    | 7  | CK260265    |
| 5          | 557.8 | 51.7  | 907    | 7  | CK266902    |
| 6          | 529.8 | 49.1  | 749    | 5  | BQ514677    |
| 7          | 523.4 | 48.5  | 794    | 4  | BM407979    |
| 8          | 521   | 48.3  | 781    | 4  | BG126704    |
| 9          | 515.2 | 47.7  | 784    | 4  | BI434774    |
| 10         | 512   | 47.5  | 742    | 7  | CK266903    |
| 11         | 443.6 | 41.1  | 639    | 5  | BQ111509    |
| 12         | 429   | 39.8  | 646    | 2  | AW621893    |
| 13         | 414.4 | 38.4  | 590    | 4  | BI928709    |
| 14         | 405.6 | 37.6  | 535    | 2  | BE433359    |
| 15         | 376.4 | 34.9  | 957    | 7  | CV267710    |
| 16         | 371.6 | 34.4  | 494    | 2  | BB924354    |
| 17         | 363.4 | 33.7  | 476    | 1  | AI772823    |
| 18         | 358.8 | 33.3  | 525    | 1  | AI775631    |
| 19         | 351.8 | 32.6  | 821    | 6  | CB290604    |
| 20         | 351   | 32.5  | 790    | 7  | CK260267    |
| 21         | 350.4 | 32.5  | 865    | 7  | CN190899    |
| 22         | 350.2 | 32.5  | 797    | 6  | CB290370    |
| 23         | 348   | 32.3  | 804    | 6  | CB293837    |
| 24         | 346.8 | 32.1  | 781    | 7  | CN190896    |

|    |       |      |     |   |          |          |           |
|----|-------|------|-----|---|----------|----------|-----------|
| 25 | 346.6 | 32.1 | 811 | 6 | CB292675 | CB292675 | UCRCS01_0 |
| 26 | 345.2 | 32.0 | 850 | 6 | CB292674 | CB292674 | UCRCS01_0 |
| 27 | 342.2 | 31.7 | 797 | 7 | CF837566 | CF837566 | UCRCS03_0 |
| 28 | 342.2 | 31.7 | 834 | 7 | CN191164 | CN191164 | UCRCS06_0 |
| 29 | 339.4 | 31.5 | 591 | 4 | BI208416 | BI208416 | EST526456 |
| 30 | 337.8 | 31.3 | 751 | 7 | CN191146 | CN191146 | UCRCS06_0 |
| 31 | 337   | 31.2 | 747 | 7 | CN189384 | CN189384 | UCRCS06_0 |
| 32 | 334.4 | 31.0 | 835 | 7 | CF830288 | CF830288 | UCRCS01_0 |
| 33 | 332.2 | 30.8 | 740 | 7 | CN191507 | CN191507 | UCRCS06_0 |
| 34 | 330.2 | 30.6 | 828 | 7 | CF838704 | CF838704 | UCRCS03_0 |
| 35 | 328.4 | 30.4 | 592 | 4 | BI204316 | BI204316 | EST522356 |
| 36 | 326.8 | 30.3 | 490 | 7 | CK270871 | CK270871 | EST716949 |
| 37 | 326.8 | 30.3 | 592 | 4 | BI204332 | BI204332 | EST522372 |
| 38 | 323.6 | 30.0 | 718 | 5 | BQ533533 | BQ533533 | QGB20812  |
| 39 | 323.4 | 30.0 | 726 | 7 | CF831499 | CF831499 | UCRCS01_0 |
| 40 | 322.4 | 29.9 | 726 | 7 | CF831872 | CF831872 | UCRCS01_0 |
| 41 | 322   | 29.8 | 546 | 2 | AW224223 | AW224223 | EST301130 |
| 42 | 319.4 | 29.6 | 793 | 7 | CF838361 | CF838361 | UCRCS03_0 |
| 43 | 317.2 | 29.4 | 712 | 5 | BQ852405 | BQ852405 | QGB17024  |
| 44 | 316.8 | 29.4 | 420 | 1 | AI772824 | AI772824 | EST253924 |
| 45 | 315   | 29.2 | 796 | 7 | CF808374 | CF808374 | peHB034XM |

ALIGNMENTS

RESULT 1  
CK269872  
LOCUS CK269872 922 bp mRNA linear EST 03-AUG-2004  
DEFINITION EST715950 potato abiotic stress cDNA library Solanum tuberosum CDNA  
clone POACP55 5' end, mRNA sequence.  
ACCESSION CK269872  
VERSION CK269872.1 GI:39826850  
KEYWORDS EST.  
SOURCE Solanum tuberosum (potato)  
ORGANISM Solanum tuberosum  
REFERENCE 1 (bases 1 to 922)  
AUTHORS Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.  
TITLE Generation of ESTs from abiotic stressed potato tissue  
JOURNAL Unpublished (2003)  
COMMENT Other ESTs: EST715951  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via http://genome.arizona.edu/orders/  
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES  
Location/Qualifiers  
1..922  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POACP55"  
/tissue\_type="abiotic stress treated leaf and root tissue"  
/lab\_host="DHI08-Tona"  
/clone\_lib="potato abiotic stress cDNA library"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: Solanum tuberosum var. Kennebec plants were  
grown from cuttings on a 16hr light/8 hr dark cycle at 25  
C for 3-4 weeks. Abiotic stress conditions were applied to  
four separate sets of plants. Set 1 involved saturation of  
the soil with 150 mM NaCl and tissues were harvested at  
following application of the salt stress (leaves: 2hr,  
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).  
Set 2 were grown under the standard conditions and then  
were water stressed by withdrawal of further watering  
applications. Drought stressed plants were harvested after  
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d





```
QY 69 GTACATGGTGCATGCCATGAGGTTGGAGTTGGTCAAGCTAAAGCCACTGCTAGAGCT 128
Db 74 GTACATGGTGCATGCCATGAGGTTGGTGTGGTATAAGCTAAAGCCATTGTTAGAGGCT 133
QY 129 GCAGGCATTAAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCACTGATTTGAGAAAAATA 188
Db 134 GCGGGCCACACGGTCACTGCCCTTGACATGGCCGCTCTGGCACTGATTTGAGAAAAATA 193
QY 189 GAGGAGCTTCGCACACTTTATGATTTATCTTTGCCATTTGATGAGTTGATGGAATCTCTT 248
Db 194 GAGGAATTTGCACACTTTGATTTATACCGTGCCATTTGATGAGTTATGGAATGCCCTT 253
QY 249 TCAGCAGATGAGAAGGTTATATTAGTGGGCATAGTCTTGGTGGTATGAATTTGGGACTT 308
Db 254 CCACAAGAGAGAAGGTCATAGTGGGCAATAGTTATGGTGTATGAATTTGGCACTT 313
QY 309 GCTATGGAAAAGTATCCACAAGATCTATGCTGCTGTTTCTTGCTGTTTCTGCTGTTTCATGCTT 368
Db 314 GCTATGGAAAATACCCAAAAAGATCTTTGTTGCTGTTTACTTGTACTGCTATTTCTGCTT 373
QY 369 GATTTCTGTTCCAACTCTCTCTTTGTTTGGACAGTATAATGAGCGGAGCCGAGCGAG 428
Db 374 GATTTCTACTCACATGCTCTCTCTGTTTGGATAAGTACTTGGAGACGACACAAGAG 433
QY 429 AATGTTGTTGATCTCAGTTTATTTACCATATGTTTCCCTGGAAGAGCCACTGACATCATG 488
Db 434 GATTTGCTTGACACCAATTTGTATCATATGTTTACCCCTGAAGAGCCTCGGACATCATG 493
QY 489 TTTTGTGGCCAAAGTTCTTGGCTCACAGCTCTACAGCTATGCTCTCTGAGGATCTT 548
Db 494 CTTTTTGGGCCCAAGTTCTTGGCTCAAAAACCTTTATCAGTTATGCTCTCTGAGGATCTT 553
QY 549 GCATTTAGCATCATTTGTTGACACCAAGCTCTTCTGTTTATGGAAGCACTATCGAAGGCC 608
Db 554 TCATTTAGCATCTGTTGTTGAGACCAACTCTCTGTTTATCGAAGATCTATCAAAAGTG 613
QY 609 AAGTATTTACAGATGAACGGTTTGGATCAGTGAAGAGAGTTTACATTTGTGCTAGT 668
Db 614 AAGTATTTACGCGACGAGGTTTGGATCAGTGAAGAGAGTTTATATATCTGCACTGAG 673
QY 669 GATAAAGCATACCAAGAGATTTCCAGCGATGCGAAATTTGACAACTTGG---TGTCACT 725
Db 674 GATAAAGTATCAACAAAAGAAATTCAGCAATGCGAAATTTGACAACTAATGGAACCGTCACT 733
QY 726 GAAGCATAGAGATTAAGGTGCTGATCATGCGCAATGCTATGCGAGCCCAAAACTT 785
Db 734 GAAGCAAGAGATCAAAAGTACTGATCATATGCGCAATGCTATGTAAGCCCAACACTT 793
QY 786 TGCGCCTCTCTCTTGGAATTTGCCATAAATCAAACTGATCTTACATTTATGTTCTGTC 845
Db 794 TGTGCGACTCTCTTGAGATTTGCCATTAATACAAATGATC-CATCTCTACATCTTCTAC 852
QY 846 TCATGTCAAGATTTTCACTGATGCTGTAATTT 878
Db 853 TCTTTTGATAAATAGTAGTTGTTGCTTTCACTT 885

RESULT 3
CK270870
LOCUS
DEFINITION
EST716948 potato abiotic stress cDNA library Solanum tuberosum cDNA
clone POACV79 5' end, mRNA sequence.
CK270870
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Solanum tuberosum (potato)
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE
1 (bases 1 to 914)
AUTHORS
Buell,C.R., Hart,A., Ziemann,V., Karamycheva,S.A. and Baker,B.
TITLE
Generation of ESTs from abiotic stressed potato tissue
```

JOURNAL  
COMMENT

Unpublished (2003)  
Other ESTs: EST716949  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via <http://genome.arizona.edu/orders/>.  
Seq primer: ANT TAG GTG ACA CTA TAG.

## FEATURES

source

1..914  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POACV79"  
/tissue\_type="abiotic stress treated leaf and root tissue"  
/lab\_host="DH10B-TonA"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: Solanum tuberosum var. Kennebec plants were  
grown from cuttings on a 16hr light/8 hr dark cycle at 25  
C for 3-4 weeks. Abiotic stress conditions were applied to  
four separate sets of plants. Set 1 involved saturation of  
the soil with 150 mM NaCl and tissues were harvested at  
following application of the salt stress (leaves: 2hr,  
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, 1 d,  
and 2d). Set 2 were grown under the standard conditions and then  
were water stressed by withdrawal of further watering  
applications. Drought stressed plants were harvested after  
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
and 5d). Set 3 were grown under the standard conditions  
and then were cold stressed by placement at 4 C. Cold  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d. Set 4 were grown under the standard conditions and  
then were heat stressed by placement at 35 C. Heat  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d and 4d and heat-stressed roots were harvested at 6 hr,  
12 hr, 1 d, and 4d. RNA was isolated from all tissues and  
equal RNA from each tissue and stress was pooled to  
construct the cDNA library. RNA sample."

## ORIGIN

Query Match 52.1%; Score 562; DB 7; Length 914;  
Best Local Similarity 80.2%; Pred. No. 4.3e-134;  
Matches 685; Conservative 0; Mismatches 165; Indels 4; Gaps 2;  
QY 9 AAAGAAAAAGAACTAACAGGCGATAAATTCAAATGAAGAGGAAAAACACTTTGTTTAA 68  
Db 51 AAAGAAAGAAAAAAACACAGCCAAAAATGGAGGTTATGAGAAACACTTTGTTTGG 110  
QY 69 GTACATGGTGCATGCGCATGGAGTTGGAGTTGGTCAAGCTAAAGCCACTGCTAGAGCT 128  
Db 111 GTACATGGTGCATGCGCATGGAGTTGGTGTGGTATAAGCTAAAGCCATTGTTAGAGCT 170  
QY 129 GCAGGCCATAGGTTACAGCCCTTCATTTAGCAGCTCTCGGCACCTGATTTGAGAAAAATA 188  
Db 171 GCGGCCCAACGCTCACTGCCCTTGACATGGCGGCTCTGGCATTTGATTTGAGAAAAATA 230  
QY 189 GAGGAGCTTCGCACACTTTATGATTTATCTTTGCCATTTGATGAGTTGATGGAATCTCTT 248  
Db 231 GAGGAATTTGACACTTTGTTGATTTATACCGTGCCATTTGATGAGTTATGGAATGCCCTT 290  
QY 249 TCAGCAGATGAGAAGGTTATATTAGTGGGCAATAGTCTTTGGTGGTATGAATTTGGGACTT 308  
Db 291 CCACAAGAGGAGAAGGTCATAGTGGGCAATAGTTATGGTGGTATGAATTTGGCACTT 350  
QY 309 GCTATGGAAAAGTATCCACAAGATCTATGCTGCTGTTTCTTGCTGCTCTTTCATGCTT 368  
Db 351 GCTATGGAAAATATACCCAAAAAGATCTTTGTTGCTGTTTACTTGCATGCTATTTCTGCTT 410  
QY 369 GATTTCTGTTCCAACTCTCTCTTTGTTTGGACAGTATAATGAGCGGAGCCGAGCGAG 428

```

Db      411  GATTCTACTACATGCTCCTATGTTTGGTAAGTACTTGGAGACACACAAAAGAG 470
QY      429  AATGGTTGGATPACTCAGTTTTTACCATATGGTTCCCTCGAAGACCACTGACATCCATG 488
Db      471  GATTGGCTTGACACCAATTTGTATCATATGTTACCCCTGAAGAGCCTCGGACATCCATG 530
QY      489  TTTTGTGGCCAAAGTTCTTGCTGCTACAAAGCTCTACCAGCTTACCACTATGCTCTCCTGAGGATCTT 548
Db      531  CTTTTTGGGCCCAAGTTCTTGCTCAAAAACCTTTATCAGTTATGCTCCCTCGAGGATCTT 590
QY      549  GCATTAGCATCATCTTGTGTGAGACCAAGCTCTTTGTTTATGGAAGACCTATCGAAGGCC 608
Db      591  TCATTAGCATCTGCTGTGTGAGACCAACCTCTCTGTTTATCGAAGATCTATCAAAAGTG 650
QY      609  AAGTATTTTACAGATGAACGGTTTGGATCAGTGAAGAGAGTTTACATTGTGTGCACTGAG 668
Db      651  AAGTATTTTACAGGACGAGGGTTTGGATCAGTGAAGAAAGTTTATATATCTGCACTGAG 710
QY      669  GATAAAGGATACCGAAGAAATTCAGCGATGGCAAAATGCAACAATTGG---TGTCACT 725
Db      711  GATAAAGTATCACCAAAAGAAATTCAGCAATGGCAAAATGCAACAATGGAAACCGTCACTC 770
QY      726  GAAGCAATAGATTAAGAGTCTGATCAGATGGAAGTCTATGCGAGCCCAAAAACCTT 785
Db      771  GAAGCAAGGAGATCAAGAGTACTGATCATATGCGCAATGCTATGTAAGCCCAACGACTT 830
QY      786  TCGCGCTCTCTCTTGGAAATTCGCCATAAATCAAACTGATCTCTACATTATGTTCTGTC 845
Db      831  TGTGCACTCTCTTGGAGATTGCCATAAATCAATTGATC-CATCTCTACATCTTCTTAC 889
QY      846  TCATGTCAAGATTT 859
Db      890  TCTTTTGAATAATT 903

```

```

RESULT 4
LOCUS   CK260265
DEFINITION   EST706343 potato abiotic stress cDNA library Solanum tuberosum cDNA
clone POAB307 5' end, mRNA sequence.
ACCESSION   CK260265
VERSION     CK260265.1 GI:39817243
SOURCE      Eukarya; Solanum tuberosum (potato)
ORGANISM   Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; famlids; Solanales; Solanaceae; Solanum.
Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
Generation of ESTs from abiotic stressed potato tissue
Unpublished (2003)
Other_ESTs: EST706342 EST706344 EST706345
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: ATT TAG GIG ACA CTA TAG.
Location/Qualifiers
1..966
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POAB307"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="PH10B-TonA"
/clone_lib="potato abiotic stress cDNA library"
/notes="Vector: pCMVSPORT5.1; Site.1: EcoRI; Site.2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were

```

## FEATURES

source

grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

## ORIGIN

```

Query Match      52.1%; Score 562; DB 7; Length 966;
Best Local Similarity 80.2%; Pred. No. 4.4e-134;
Matches 685; Conservative 0; Mismatches 165; Indels 4; Gaps 2;

QY      9  AAGAAAAAGAAATAACAAGGCATAAAATTCAAATGAAGGAAGGAAACAACCTTTGTTTAA 68
Db      15  AAGAAAAAGAAATAACAAGGCATAAAATTCAAATGAAGGAAGGAAACAACCTTTGTTTAA 74
QY      69  GTACATGCTGTCATGCCATGGAGTTGGAGTTGGTACAGCTTAAGGCCACTCTAGAACCT 128
Db      75  GTACATGCTGTCATGCCATGGAGTTGGTGGTATGAAGCTTAAGGCCACTCTTAGAGGCT 134
QY      129  GCAGGCCATAAGGTTACAGCCCTTGATTAGCAGCTCTGGCAGCTGATTTGAGAAAAATA 188
Db      135  GCGGCCACACGGTCACCTGCTGACATGCCGCTCTGGCATGTGATTGAGAAAAATA 194
QY      189  GAGGAGCTTCGCACACTTTATGATATATCTTTGCCATTTGAGGAGTTGATGGAATCTCTT 248
Db      195  GAGGAACTTTGCACACTTTGTTGATTATACCTGCCATTGATGGAGTTTATGGAATGCCTT 254
QY      249  TCAGCAGATGAGAAGTTATATTAGTGGGCGATAGCTCTGGTGGTATGAATTTGGGACTT 308
Db      255  CCAAGAGGAGGAAGGTCACTACTAGTGGGCGATAGTTATGGTGGTATGAATTTGGCATT 314
QY      309  GCTATGAAAAAGTATCCACAAAAGATCTATGCTGCTGTTTCTTGGCTGCTTTTCATGCT 368
Db      315  GCTATGAAAAAGTATCCCAAAAAGATCTTGTGCTGTTTACTTGTACTGCTATTTGCTCT 374
QY      369  GATTCCTGTTCAAACTCTCTCTCTTTGTTGAAACAGTATAATAGCGGAGCGCAGCCGAG 428
Db      375  GATTCCTACTCATGCTCTCTCTATGTTTGGATAGTACTTTGGAGACGACACAAAAGAG 434
QY      429  AATTGGTTGGATCTACGTTTTTACCATATGGTTCCCTCGAAGAGCCACTGCATCCATG 488
Db      435  GATTGGCTTGACACCCAAATTTGTATCATATGGTACCCCTCGAAGAGCCCTCGGACATCCATG 494
QY      489  TTTTGTGGCCAAAGTTCTTGGCTCACAAGCTCTACCAAGCTATGCTCTCTGAGGATCTT 548
Db      495  CTTTTTGGGCCCAAGTTCTTGGCTCAAAAACCTTATCAGTTATGCTCTCTCTGAGGATCTT 554
QY      549  GCATTAGCATCATCTGCTGAGACCAAGCTCTTTGTTTATGGAAGACCTATTCGAAGGCC 608
Db      555  TCATTAGCATCTGCTGCTGAGACCAACCTCTCTGTTTATCGAAGATCTATCAAAAGTG 614
QY      609  AAGTATTTTACAGATGAACGGTTTGGATCAGTGAAGAGGTTTACATTTGTGCACTGAG 668
Db      615  AAGTATTTTACCGGACGAGGGTTTGGATCAGTGAAGAAAGTTTATATATCTGCACTGAG 674
QY      669  GATAAAGGATACCAAGAAATTCAGCGATGGCAAAATGCAACAATTGG---TGTCACT 725

```



|                           |    |     |  |             |
|---------------------------|----|-----|--|-------------|
|                           | Qy | 541 | AGCATCTTGGATTAGCATCATCATTTGCTGTGAGACCAAGCTCTTTCGTTTTATCGAAGACCTTAT | 600         |
|                           | Db | 481 | AGATCTTTTCATTAGCATCTGCTGTTGTGTGAGACCACCTCTCTCTTTATCGAAGATCTTAT     | 540         |
|                           | Qy | 601 | CGAAGGCCAAAGTATTTCCACAGATGAACGGTTTTGGATCAGTGGAAGAGAGTTTTACATTTGTT  | 660         |
|                           | Db | 541 | CAGAAGTGAAGTATTTTCACGCGACGAAGGGTTTGGATCAGTGAAGAAAAGTTTTATAATCT     | 600         |
|                           | Qy | 661 | GCACTGAGGATAAAGGCATACCGAGAAGAATCCAGCGATGCGCAATTTGACAACATTTGG--     | 718         |
|                           | Db | 601 | GCACTGAGGATAAAGTATATCACCAAAGAAATTCACGCAATGCGCAATTTGACAACATTTGAA    | 660         |
|                           | Qy | 719 | TGTCCTACTGAAGCAATAGAGATTAAAGGTGCTGTATCAGTGGCAATGCTATGCGAGCCCC      | 777         |
|                           | Db | 661 | CGTTCATCGAAGCAAGGAGATCAAGGTACTGATCATATGCGCAATGCTATGTAAGCCCA        | 720         |
|                           | Qy | 778 | AAAAAATTTGCGCTCTCTCTTTGAAATT                                       | 806         |
|                           | Db | 721 | AACGACTTTTTCGCACTCTCTTTGAGATT                                      | 749         |
| RESULT 7                  |    |     |  |             |
| BM407979                  |    |     |  |             |
| LOCUS                     |    |     |  |             |
| DEFINITION                |    |     | 794 bp mRNA linear EST 10-MAR-2003                                 |             |
|                           |    |     | EST582306 potato roots Solanum tuberosum cDNA clone cPRO32J8 5'    |             |
|                           |    |     | end, mRNA sequence.  |             |
| ACCESSION                 |    |     | BM407979   |             |
| VERSION                   |    |     | BM407979.1   | GI:18259609 |
| KEYWORDS                  |    |     | EST.   |             |
| SOURCE                    |    |     | Solanum tuberosum (potato)   |             |
| ORGANISM                  |    |     | Solanum tuberosum  |             |
|                           |    |     | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; |             |
|                           |    |     | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;       |             |
|                           |    |     | asterids; lamiales; Solanales; Solanaceae; Solanum.                |             |
| REFERENCE                 |    |     | 1 (bases 1 to 794)   |             |
| AUTHORS                   |    |     | van der Hoeven,R., Sun,H., Karamycheva,S.A., Tsai,J., Van Aken,S., |             |
|                           |    |     | Utterback,T., Chienming,A., Bougri,O., Buell,C.R., Renning,C.,     |             |
|                           |    |     | Tankley,S. and Baker,B.  |             |
|                           |    |     | Generation of ESTs from potato roots                               |             |
| TITLE                     |    |     | Unpublished (2001)   |             |
| JOURNAL                   |    |     | Contact: Robin Buell   |             |
| COMMENT                   |    |     | The Institute for Genomic Research                                 |             |
|                           |    |     | 9712 Medical Center Dr, Rockville, MD 20850, USA                   |             |
|                           |    |     | Email: potato-array@igr.org  |             |
|                           |    |     | This clone can be obtained from the University of Arizona Genomics |             |
|                           |    |     | Institute. Orders can be made through URL:                         |             |
|                           |    |     | http://genome.arizona.edu/orders/                                  |             |
|                           |    |     | Seq primer: T3   |             |
| FEATURES                  |    |     | Location/Qualifiers  |             |
| source                    |    |     | 1..794   |             |
|                           |    |     | /organism="Solanum tuberosum"                                      |             |
|                           |    |     | /mol_type="mRNA"   |             |
|                           |    |     | /cultivar="Kennebec"   |             |
|                           |    |     | /db_xref="taxon:4113"  |             |
|                           |    |     | /clone="cPRO32J8"  |             |
|                           |    |     | /tissue_type="roots"   |             |
|                           |    |     | /dev_stage="in vitro grown stem cuttings"                          |             |
|                           |    |     | /lab_host="SOLR"   |             |
|                           |    |     | /clone_lib="potato roots"  |             |
|                           |    |     | /note=vector: plasmid SK(-); Site 1: EcoRI; Site 2:                |             |
|                           |    |     | XhoI; supplier: Cornell University, Tankley lab;                   |             |
|                           |    |     | sequencing; The Institute for Genomic Research. Roots were         |             |
|                           |    |     | isolated from in vitro grown stem cuttings on CM medium.           |             |
|                           |    |     | Roots were isolated two weeks after placing the stem               |             |
|                           |    |     | cuttings from in vitro grown plants on medium."                    |             |
| ORIGIN                    |    |     |  |             |
| Query Match               |    |     | 48.5%; Score 523.4; DB 4; Length 794;                              |             |
| Best Local Similarity     |    |     | 80.7%; Pred No. 4e-124;  |             |
| Matches 624; Conservative |    |     | 0; Mismatches 146; Indels 3; Gaps 1                                |             |
| Qy                        |    |     | 9 AAAGAAAGAACTAACAAAGSCATATAAATTCAAATGAAGGAAGGAAAAACATTTGTTT       | 68          |

Db 22 AAGAAAGAAAAAACAACAAGCAAAAAAATGGAGTTATGAAGAAACACTTTGTTTG 81  
Qy 69 GTACATGGTGCATGCCATGGAGTTGGAGTTGGTCAAGCTAAAGCCACTGCTAGAAGCT 128  
Db 82 GTACATGGTGCATGCCATGGAGTTGGTGGTATGAAGCTAAAGCCATTTAGAGGCT 141  
Qy 129 GCAGGCCATTAAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCACTGAATTTGAGAAAAATA 188  
Db 142 GCGGGCCACACGGTCACTGCCCTTGACATGGCGGCTCTGGCATTTGAGAAAAATA 201  
Qy 189 GAGGACCTTCGACACTTATGATATATCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 248  
Db 202 GAGGAATTTGACACTTGTGATATACCGTGCCATGATGGAGTTTATGAATGCGCT 261  
Qy 249 TCAGCAGATGAGAGGTTATATTAGTGGGCGCATAGTCTTGGTGGTATGAATTTGGGACTT 308  
Db 262 CCACAGAGAGAGGTCATCTAGTGGGCGCATAGTTATGGTGGTATGAATTTGGGACTT 321  
Qy 309 GCTATGGAAGATATCCACAAAGATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 368  
Db 322 GCTATGGAAGATATCCACAAAGATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 381  
Qy 369 GATTCGTTTCACAACTCTCTCTTTGTTTGGAAACAGTATAAGAGGGAGCGCCAGCGAG 428  
Db 382 GATTCCTACTCACATGCTCTCTCTATGTTTGGATAAGTACTTGGAGACGCAACAAAAAGAG 441  
Qy 429 AATTGGTGGTACTCAGTTTATACCATATGTTTCCCTGGAAGCCACTGACATCCATG 488  
Db 442 GATTTGGCTTGACACCCAAATTTGTATCATATGGTACCCCTGGAAGGCTCGGACATCCATG 501  
Qy 489 TTTTGGGCCCCAAGTCTTGGCTCAAGCTCTACCACTATGCTCTCTGAGGATCTT 548  
Db 502 CTTTGGGCCCCAAGTCTTGGCTCAAAACTTATCACTATGCTCTGAGGATCTT 561  
Qy 549 GCATTAGCATCATATGTTGGTAGACCAAGCTCTTTGTTTATGGAAGACCTATCGAAGGCC 608  
Db 562 TCATTAGCATCGTGGTGGTAGACCAAGCTCTCTGTTTATGGAAGATCTATCAAAAGGTG 621  
Qy 609 AAGTATTTACAGATGAACGGTTTGGATCAGTGAAGAGAGTTTACATGTTGGTCACTGAG 668  
Db 622 AAGTATTTACGAGACGAGGTTTGGATCAGTGAAGAGAGTTTATATATCTGCACTGAG 681  
Qy 669 GATAAGGATACCAAGAAATTCACGAGATGCGCAAAATTGACAACTTTGG---TGTCACT 725  
Db 682 GATAAGTATACCAAAAGATTCACGCAATGCGCAAAATTGACCAACCATGGGACCGTCATC 741  
Qy 726 GAAGCAATAGATTAAGGTTGCTGATCATGAGGCAATGCTATGCGAGGCCCA 778  
Db 742 GAAGCAAGGAGATCAAGGTTACTGATCATATGCGCAATGCTATGTAAGGCCAA 794

## RESULT 8

BG126704

LOCUS

DEFINITION EST472350 tomato shoot/meristem Lycopersicon esculentum cDNA clone cTOF13K13 5' sequence, mRNA sequence.

ACCESSION BG126704

VERSION BG126704.1

KEYWORDS GI:12626892

SOURCE EST.

ORGANISM

Lycopersicon esculentum (tomato)  
Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; famids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2001)

Contact: CUGI

Clemson University

Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

## FEATURES

source

1. 781

Location/Qualifiers  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cTOF13K13"  
/tissue\_type="shoot/meristem"  
/dev\_stage="developing shoots from 4-6wks old plants"  
/lab\_host="SOUR"  
/clone\_lib="tomato shoot/meristem"  
/note="vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
XhoI; Small expanding leaves from the growing tip were  
taken from greenhouse plants (4-6wks old TA496). Tissue  
was immediately frozen in liquid nitrogen."

## ORIGIN

Query Match 48.3%; Score 521; DB 4; Length 781;  
Best Local Similarity 82.6%; Pred. No. 1.7e-123;  
Matches 609; Conservative 0; Mismatches 125; Indels 3; Gaps 1;  
Qy 42 ATGAAGGAAGAAACACTTTGTTTGTAGTACATGGTGCATGCCATGGAGTTGGATTGG 101  
Db 45 AAGAAACAAGGAAGACACTTTGTTTGTGTACATGGTGCATGCCATGGAGTTGGATTGG 104  
Qy 102 TACAGCTAAAGCCACTGCTAGAACTGCAGGCCATAGGTTACAGCCCTTGATTAGCA 161  
Db 105 TACAACTAAAGCCCTTCTAGAGGTTGCAGGCCACAAGGTCACAACTCTTGATTGGCA 164  
Qy 162 GCTTCTGGCACTGATTTTGAGAAAAATAGAGAGCTTCCACACTTTATGATTATATCTTTG 221  
Db 165 GCCTCTGTTATGATTTGAGAAAAATAGAGCAACTTCACACATTTTCATGATTACACTTTG 224  
Qy 222 CCATTGATGAGGATGATGGAATCTCTTTCAGCAGATGAGAAAGTTATATATTAGTGGGGCAT 281  
Db 225 CCATTGATGGAATTTGATGGAATCTCTTCCAAAGAGGAGAAAGTCATACTAGTTGGACAT 284  
Qy 282 AGTCTGGTGTGATGAATTTGGGACTTCTATGAAAAGATATCCAAAAAGATCTATGCT 341  
Db 285 AGTCTGGTGTGATGAATTTGGGACTTCTATGAAAAGATATCCAAAAAGATCTATGCT 344  
Qy 342 GCTGTTTTCTTTGGCTGCTTTCATGCTGATTTCTGTTCACAACTCTCTCTTTTGGAA 401  
Db 345 GCTGTTTTCTTTGGCTGCTTTCATGCTGATTTCTATTCATCTGCTCTTGTCTTGGAT 404  
Qy 402 CAGTATATAGCGGACCGCAGAGAAATTTGGTGGATCTCAGTATCTATGTTTACCATATGCT 461  
Db 405 CAGTATTTTGAGAGGATGCAAAACAATGAATTTGGTGGACACCCCAATTTGTATCATATGCT 464  
Qy 462 TCCCTGGAAGAGCCACTGACATCCATGTTTTCGCCCCAAAGTTCTTGGCTCACAAGCTC 521  
Db 465 TCCCATGAAGAGCCCTTGGCCATTCATATTTTTCGCCCCAAAGTTCTTGGCTTACAACTCT 524  
Qy 522 TACCAGCTATCTCTCTGAGGATCTTGCATTTAGCATCATATTTGGTGGAGACCAGCTCT 581  
Db 525 TATCAATTATGCTCTCTGAGGATGTTGATGTTGATCATCATTTGGGAGAGCAAGTTCT 584  
Qy 582 TTGTTTATGGAAGACCTATCGAAGCCCAAGTATTTTCAGATGAACGGTTTGGATCAGTG 641  
Db 585 TTAATCTTAGAAGATTTGTCAAAATCAAAATTTTGAAGATGAAGGATATGATGATCAGTG 644  
Qy 642 AAGAGAGTTTACATTTGTCAGCTGAGGATTAAGGCAATACCAGAGAAATTCACGCGATGG 701  
Db 645 AAGAAAGTTTATATAGTATGACAGATGAATAAATCTCTTACCACAAAGAGTTTCAAAATGG 704  
Qy 702 CAATTTGACAAACAT---TGGTGTCTCACTGAAGCAATAGAGATTAAGAGTGTGATCACATG 758  
Db 705 CAATTTGACAAACATTAATAGCATATAGAAACAAGAAATTTGAAGTGTGATCATATG 764  
Qy 759 GCAATGCTATCGGAGCC 775  
Db 765 GCAATGCTAAGTATGCC 781





/note=vector; pCWSport6.1; Site\_1: EcoRI; Site\_2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to C for separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at the following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots:3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

## ORIGIN

| Query Match           | 47.5%;          | Score 512;  | DB 7;     | Length 742; |
|-----------------------|-----------------|---|-----------|-------------|
| Best Local Similarity | 82.4%;          | Prod. No. 3.5e-121;   |           |             |
| Matches 600;          | Conservative 0; | Mismatches 125;   | Indels 3; | Gaps 1      |
| Qy                    | 102             | TACAAAGCTAAAGCCACTGCTAGAACCTCGAGGCCATAGGTTACAGCCCTTGATTTAGCA        | 161       |             |
| Db                    | 742             | TATAAGCTAAAGCCCAATTGTTAGAGGCTCGCGGCCACACGGTCACCTGCGCCTTGACATGGCC    | 683       |             |
| Qy                    | 162             | GCTTCTGGCACTGATTTGAGAAAAATAGAGGAGCTTCGCACACTTTATGATTATATACTTTG      | 221       |             |
| Db                    | 682             | GCTTCTGGCAATTGTTGAGAAAAATAGAGGAACTTTGGCACACTTGTGTGATTATACCGTG       | 623       |             |
| Qy                    | 222             | CCATTGATGAGTTGATGGAAATCTCTTTTCCAGCAGATGAGAAGTTATATTAGTGGGGCAT       | 281       |             |
| Db                    | 622             | CCATTGATGGAGTTTATGGAAATGCCTTTCCAAAGAGGAGNAGGTCATACTAGTGGGGCAT       | 563       |             |
| Qy                    | 282             | AGTCTTGGTGTATGAATTTTGGGACTTCGTATGGAAAAAGTATCCACAAAGATCTATGCT        | 341       |             |
| Db                    | 562             | AGTTATGGTGTGTGAATTTGGCACTTCGTATGGAAAAATACCCAAAAAGATCTTTGTT          | 503       |             |
| Qy                    | 342             | GCTGTTTTCTTGGCTGCTTTTCATGCTCTGATTTCTGTTTCCAACTCTCTCTTCTGTTTTGGAA    | 401       |             |
| Db                    | 502             | GCTGTTTACTTTGACTGCTTATTTCTGCTCTGATTTCTACATGTCCTCTCTATGTTTGGAT       | 443       |             |
| Qy                    | 402             | CAGTATATAGCGGACGCCAGCCAGCAAGTAATGGTTGGATCTACTCAGTTTTTACCATATGCT     | 461       |             |
| Db                    | 442             | AAGTACTTTGGAGACGACAAACAAAGAGGATTTGGCTTGCACCCAAATTTGTATCATATGCT      | 383       |             |
| Qy                    | 462             | TCCCTCTGAAGAGCCACTGACATGCCATGCCATGTTTTTTTGGCCCAAAGTTCTTGCTCTCAAGCTC | 521       |             |
| Db                    | 382             | ACCCCTGAAGAGCCCTCGGACATCCATGCTTTTGTGGGCCCAAGTTCTTGCTCTCAAAAACCTT    | 323       |             |
| Qy                    | 522             | TACCAGCTATGCTCTCCCTGAGGATCTTCGATTTAGCATCATCATTTGGTGAGACCAAGCTCT     | 581       |             |
| Db                    | 322             | TATCAGTTTATGCCCTCTCTGAGGATCTTTTTCATTTAGCATCGTGGTGGTGAACCAACCTCT     | 263       |             |
| Qy                    | 582             | TTGTTTTATGGAAGACCTATCGAAGGCCAAGTATTTTCCAGATGAACGGTTTGGATTCAGTG      | 641       |             |
| Db                    | 262             | CTGTTTATCGAAGATCTATCAAAGTGAAGTATTTTCCGACGAAGGGTTTGGATTCAGTG         | 203       |             |
| Qy                    | 642             | AAGAGAGTTTACATTTGTGTGCACTGAGGATAAAGGCATACCAGAGAAGATTCACAGCGATGG     | 701       |             |
| Db                    | 202             | AAGAAAGTTTATATAATCTGCACCTGAGGATAAAGTATCACCCAAAAGAAATCCAGCAATGG      | 143       |             |
| Qy                    | 702             | CAAAATGCAACATTTGG---TGTCACTGAAGCAATAGAGATTAAAGGTCCTGATCACATG        | 758       |             |
| Db                    | 142             | CAAAATGCAACAAATGGAACCGTCAATCGAAGCAAGAGGAGATCAAAAGGTACTGATCATATG     | 83        |             |

|    |     |  |     |
|----|-----|--|-----|
| QY | 759 | GCAATGCTATCGGAGCCCCAAAACCTTTTGGCCCTCTCTCTCTTGGAAATTTGCCCATAAATAC | 818 |
| Db | 82  | GCATGCTATGTAAGAGCCCAACGACTTTTGTGCACTCTCTTGGAGATTGCCCATAAATAC     | 23  |
| QY | 819 | AACTGATC   | 826 |
| Db | 22  | AATTGATC   | 15  |

RESULT 11

| LOCUS    | DEFINITION |
|----------|------------|
| BQ111503 |            |

BQ111509 639 bp mRNA linear EST597085 mixed potato tissues Solanum tuberosum cDNA clone STMCB2115' end, mRNA sequence. EST 07-MAR-2003

ACCESSION  
VERSION  
KEYWORDS

BQ111509  
BQ111509.2 GI:21914979  
EST.

**NETWORKS**  
**SOURCE**  
**ORGANISM**

Solanum tuberosum (potato)  
Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum.  
 REFERENCE 1 (bases 1 to 639)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Solanum.  
1 (bases 1 to 639)  
Buell, C.R., Hart, A., Baker, B., Tankley, S., Fry, W., Smart, C.,  
Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and  
Karamycheva, S.A.

| TITLE  | JOURNAL                         | COMMENT                           |
|--|---------------------------------|-----------------------------------|
| 1. The Role of the Teacher in the Classroom                  | Journal of Educational Research | 1980, Vol. 83, No. 1, pp. 1-10    |
| 2. The Impact of Technology on Education                     | Journal of Educational Research | 1980, Vol. 83, No. 2, pp. 11-20   |
| 3. The Importance of Parental Involvement                    | Journal of Educational Research | 1980, Vol. 83, No. 3, pp. 21-30   |
| 4. The Effect of Teacher Expectations on Student Achievement | Journal of Educational Research | 1980, Vol. 83, No. 4, pp. 31-40   |
| 5. The Role of the School in the Community                   | Journal of Educational Research | 1980, Vol. 83, No. 5, pp. 41-50   |
| 6. The Impact of Teacher Education on Student Achievement    | Journal of Educational Research | 1980, Vol. 83, No. 6, pp. 51-60   |
| 7. The Importance of Teacher Professionalism                 | Journal of Educational Research | 1980, Vol. 83, No. 7, pp. 61-70   |
| 8. The Effect of Teacher Salary on Student Achievement       | Journal of Educational Research | 1980, Vol. 83, No. 8, pp. 71-80   |
| 9. The Role of the School in the Community                   | Journal of Educational Research | 1980, Vol. 83, No. 9, pp. 81-90   |
| 10. The Impact of Teacher Education on Student Achievement   | Journal of Educational Research | 1980, Vol. 83, No. 10, pp. 91-100 |

Generation of a set of potato cDNA clones for microarray analyses  
Unpublished (2002)  
On Apr 17, 2002 this sequence version replaced qi:20163471.

Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: [potato-array@igr.org](mailto:potato-array@igr.org)  
This clone can be obtained from the University of Arizona Genomics  
Institute. Orders can be made through URL:  
<http://genome.arizona.edu/orders/>

## FEATURES

**source**

Location/Qualifiers  
1. .639

/organism="Solanum tuberosum"

```

/mo1 type="mrna"

```

/cultivar="Kennebec or Binjite"

```
/db_xref="taxon:4113"
```

```

/clone="STMCB21"

```

```
/tissue type="mixed tissues"
```

```
/lab host="SOLR"
```

```
/clone lib="mixed potato tissues"
```

```
/note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Combination of untreated and Phytophthora
infans-treated libraries of stolons, leaves, leaflets,
axillary buds of stem explants, petioles, germinating
aves, tubers, or roots."
```

## ORIGIN

|                       |              |                     |                |                  |
|-----------------------|--------------|---------------------|----------------|------------------|
| Query Match           | 41.1%;       | Score 443.6;        | DB 5;          | Length 639;      |
| Best Local Similarity | 83.6%;       | Pred. No. 1.7e-103; |                |                  |
| Matches 503:          | Conservative | 0;                  | Mismatches 99; | Indels 0; Gaps 0 |

114 CCACTGCTAGAGTGCAGGCCATAGGTTACAGCCCTTGATTAGCAGTTCCTGGCACT 173  
|||||  
|||||  
|||||  
95 CCAATTGTTAGAGTGTCCGACCAACGGTCACCTGCCCTTGACATGGCCGCTCTCGCATTT 154  
|||||  
|||||  
|||||

174 GATTTGAGAAAAATAGAGGAGCTTGGCACATTTATGATTATATCTTGGCATTGATGGAG 233  
|||||  
155 GATTTGAGAAAAATAGAGAACTTTGGCACATTTGTTGATTATACCGTGCATTGATGGAG 214  
|||||

234 TTGATGGGAATCTCTTTCAGCAGATGAGAAGTTATAATTAGTGGGGCATACTCTTGCGTGGT 293  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
215 TTTTATCGAATGCCCTTCCAAAGACGACGAGGCATATTACTGGGGCATACTTATGCGTGGT 274

```

QY 294 ATGAATTTGGGACTTGCTATGGAAGAAAGTATCCACAAAAGATCTATGCTGCTGTTTCTTG 353
Db 275 ATGAATTTGGGACTTGCTATGGAAGAAATACCCAAAAAGATCTTTGTTGCTGTTTACTTG 334
QY 354 GGTGCTTTTCATCCCTGATTCGTTCACAACTCCTCTTTGTTTGGAAACAGTATATGAG 413
Db 335 ACTGCTATTCTGCTGATTTCTACTCACATGTCCTCTATGTTTGGATAAGTACTTGGAG 394
QY 414 CGGAGCCGACCGAGAAATTTGGATATCTCAGTTTATTTTACCATATGTTTCCCTCGAAGAG 473
Db 395 ACGACAAACAAAGAGATTTGGCTTGACACCCAAATTTGATATCATATGTTTACCCCTGAAGAG 454
QY 474 CCATGACATCATGTTTGTGTTTGGCCCAAAAGTTCTTGGCTCAAAAGCTCTACCAAGTATGC 533
Db 455 CCTCGACATCATGTTTGTGTTTGGCCCAAGTTCTTGGCTCAAAAGCTCTTATCAGTTATGC 514
QY 534 TCTCTGAGATCTTGATATGATATCATATTTGTTGGTGGACCAAGCTCTTGTGTTATGAA 593
Db 515 CCTCTGAGATCTTTTCAATAGCATGCTGTTGGTGAGACCAACCTCTCTGTTTATCGAA 574
QY 594 GACCTATCGAAGCCCAAGTATTTACAGATGAACGGTTTGGATCAGTGAAGAGAGTTTAC 653
Db 575 GATCATCAAGGTGAAGTATTTACGAGACGAAGGTTTGGATCAGTGAAGAAAGTTTAT 634
QY 654 AT 655
Db 635 AT 636

```

```

RESULT 12
AW621893 646 bp mRNA linear EST 18-MAY-2001
LOCUS EST1212691 tomato root during/after fruit set, Cornell University
DEFINITION Lycopersicon esculentum cDNA clone cLEX13L15 5', mRNA sequence.
ACCESSION AW621893
VERSION AW621893
KEYWORDS EST.
SOURCE AW621893.1 GI:7333540

```

```

ORGANISM Lycopersicon esculentum (tomato)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 646)
van der Hoeven,R.S., Garvin,D., Matern,A.L., Holt,I.E., Kochian,L.,
Liang,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and
Tanksley,S.D.
Generation of ESTs from tomato root, during and after fruit set
Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
Location/Qualifiers
1. .646
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEX13L15"
/tissue_type="root"
/clone_lib="tomato root during and after fruit-set"
/notes="vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Tanksley; Tissue supplied by Dave Garvin
(USDA-ARS, Ithaca, NY 14850)."

```

```

FEATURES
source
1. .646
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEX13L15"
/tissue_type="root"
/clone_lib="tomato root during and after fruit-set"
/notes="vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Tanksley; Tissue supplied by Dave Garvin
(USDA-ARS, Ithaca, NY 14850)."
ORIGIN
Query Match 39.8%; Score 429; DB 2; Length 646;

```

```

Best Local Similarity 79.3%; Pred. No. 9.9e-100;
Matches 545; Conservative 0; Mismatches 90; Indels 52; Gaps 1;
QY 42 ATGAAGGAAGAAACACATTTTGTATGATACATGGTGCATGCCATGGAGGTTGGAGTTGG 101
Db 12 ATGAAGGAAGAAACATTTTGTATGATACATGGTGCATGCCATGGAGGTTGGAGTTGG 71
QY 102 TACAAGCTAAAGCCACATCTGCTAGAAGCTGCAGGCCATAAGGTTACAGCCCTGATTAGCA 161
Db 72 CACAAATTTAAACACATTTGTTGGAGGGTGCAGGCCATAAGGTCACCTGCCCTTGATTTAGCA 131
QY 162 GCTTCTGCACTGATTTGAGAAAAATAGAGAGAGCTTGCCACACTTTATGATTATATCTTTG 221
Db 132 GCCTCTGGGATGATTTGAGAAAAATACACACTTTACACACTTTATGATTATACCTTTG 191
QY 222 CCATTGATGGAGTTGATGGAATCTCTTTTCAGCAGATGAGAAAGGTTATATTAGTGGGGCAT 281
Db 192 CCATTGATGGAGTTGATGGAATCTCTTTCCACAAGAGAGAAAGTCATACTAGTTGGACAT 251
QY 282 AGTCTTGGTGTATGAATTTGGGACTTGCTATGGAAGATATCCACAAAAGATCTATGCT 341
Db 252 AGTCTTGGTGTATGAATTTGGGACTTGCTATGGAAGATATCCCTAAAGAGATTTATGTT 311
QY 342 GCTGTTTTCTTGGCTGCTTTCATGCTGATTTCTTCAAACTCTCTCTTTGTTTGGAA 401
Db 312 GCTGTTTTCTTGGCTGCTTTCATGCTGATTTCTTCAATGTCATCTATGTTCTTGAAT 371
QY 402 CAGTATATAGCGGACGCCAGCCAGAAATTTGGTGGATCTCAGTATCTATGTTTACCATATGCT 461
Db 372 CAGTACTATGAACGGACGCCAGAGAAATTTGGTTAGACACTCAATTTTACCATATGCT 431
QY 462 TCCCTCTGAAGCCACATGACATCCATGTTTGGCCCAAGATTTCTGGCTCACAAAGCTC 521
Db 432 ACACCTGAAGAGCCACATCCTCATGTTTGGCCCTAAATTTTGGCTCACAAAGCTC 491
QY 522 TACCAGCTATGCTCTCTCTGAGGATCTTGCATTAGCATCATGTTGGTGAGACCAAGCTCT 581
Db 492 TACCATTATGCTCTCTG----- 510
QY 582 TTGTTATGGAAGACCTATCGAAGCCCAAGTATTTTCAGATGAACGGTTTGGATGATG 641
Db 511 -----AGATTGTCAAAGGAAAGTACTTGGTATGAAGTATATGATCGGTTG 559
QY 642 AAGAGAGTTTACATTTGTGCTGCACTGAGATAAAGGCATACCAGAGATTTCCAGCGATGG 701
Db 560 AAGAGAGTTTACATAGTATGCACTGAGCAAAAGGCATACCACAAAGATTTCAACAATGG 619
QY 702 CAATTGACAACTTTGGTGTCACTGAA 728
Db 620 CAATTGACAACTTTGGTGTGACAGAA 646

```

```

RESULT 13
BI928709 590 bp mRNA linear EST 18-OCT-2001
LOCUS EST548598 tomato flower, 3 - 8 mm buds Lycopersicon esculentum CDNA
DEFINITION clone cTOB26K14 5' end, mRNA sequence.
ACCESSION BI928709
VERSION BI928709.1 GI:16240757
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 590)
van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
Uterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)
Unpublished (2001)
Contact: CUGI
Clemson University Genomics Institute

```



Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
This clone is available through the Clemson University Genomics Institute  
Seq primer: T3.  
Location/Qualifiers  
1. 590  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CT0826K14"  
/tissue\_type="flower"  
/dev\_stage="3-8mm buds"  
/clone\_lib="tomato flower, 3 - 8 mm buds"  
/notes="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2: XhoI; supplier: Cornell University; sequencing: The Institute for Genomic Research. Flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."  
\*  
Query Match 38.4%; Score 414.4; DB 4; Length 590;  
Best Local Similarity 84.4%; Pred. No. 5.8e-96;  
Matches 466; Conservative 0; Mismatches 86; Indels 0; Gaps 0;  
QY 42 ATGAAGGAGGAAAAACATTGTTTGTAGTACATGTTGCGATGCCATGGAGTTGGAGTTGG 101  
DB 35 AAGAAAACAAGGAAGACATTGTTTGGTACATGTTGCGATGCCATGGAGTTGGTGG 94  
QY 102 TACAAGCTAAGCCACTGCTAGAGCTGCAGGCCATAAGSTTACAGCCCTTGATTAGCA 161  
DB 95 TACAACTAAGCCCCCTTAGAGTTGACAGGCCACAAGTCACAACTCTGATTGGCA 154  
QY 162 GCTTCTGGCACTGATTGAGAAAAATAGAGAGCTTCGCACACTTTATGATTATCTTTG 221  
DB 155 GCCTCTGGTATTGATTGAGAAAAATAGAGCAACTTCACACATTTTCATGATTACATTG 214  
QY 222 CATTGATGAGTTGATGAATCTCTTTCAGCAGATGAGAAGTTATATAGTGGGCAT 281  
DB 215 CCATTGATGGAATTGATGGAATCTCTTCCACAAGAGGAGAAAGTCATCTAGTTGGACAT 274  
QY 282 AGTCTTGGTGTATGAATTTGGACTTGTCTATGGAAGATATCCCAAAGATCTATGCT 341  
DB 275 AGTCTTGGTGTATGAATTTGGACTTGTCTATGGAAGATATCCCAAAGATCTATGCT 334  
QY 342 GCTGTTTTCTTGGCTGCTTTTCATGCTGATTCTGTTTCAACAACCTCTCTCTTTTGGAA 401  
DB 335 GCTGTTTTCTTGGCTGCTTTTCATGCTGATTCTATTTCACTCTCTCTCTATGCTTGGAT 394  
QY 402 CAGTATAATGAGCGGACGCCAGCGGAGAAATGGTTGGATATCTCAGTTTTTACATATGTT 461  
DB 395 CAGTATTTTGGAGGAGTCAAAACAATGAATTTGGTGGACACCCAAATTTGTATCATATGTT 454  
QY 462 TCCTCTGAAGACCACTGACATCCATGTTTTTTGGCCCAAGTCTTGGCTCACAGCTC 521  
DB 455 TCCCATGAAGAGCCTCTGCTCATCATATTTTTTGGGCCCAAGTCTTGGCTTACATCTT 514  
QY 522 TACCAGCTATGCTCTCTGAGGATCTTGCATTAGCATCATCTTGTGTAGACACAAGCTCT 581  
DB 515 TATCAATTATGCCCTCTGAGGATGTTGCAATGGTATCATCTTTGGGAGACCAAGTTCT 574  
QY 582 TTGTTTATGGAA 593  
DB 575 TTATTCCTAGAA 586  
RESULT 14  
BE433359  
LOCUS  
DEFINITION EST399888 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA  
clone cLEG13E6, mRNA sequence.  
BE433359  
BE433359.1 GI:9431202  
EST.  
Lycopersicon esculentum (tomato)  
Lycopersicon esculentum  
Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 535)  
Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,  
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Romning,C.M.,  
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and  
Tanksley,S.D.  
Generation of ESTs from tomato fruit tissue, breaker stage  
Unpublished (2000)  
Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.  
Location/Qualifiers  
1. 535  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cLEG13E6"  
/tissue\_type="Pericarp"  
/dev\_stage="breaker"  
/lab\_host="SOLR"  
/clone\_lib="tomato breaker fruit, TIGR"  
/note="Vector: pBluescriptSKmCuadapt; Site 1: EcoRI;  
Site 2: XhoI; Fruit were harvested at the breaker stage  
(first sign of lycopene accumulation on the blossom end of  
the fruit). Fruit were cut in half and the seeds and  
locules were discarded prior to freezing the pericarp."  
ORIGIN  
Query Match 37.6%; Score 405.6; DB 2; Length 535;  
Best Local Similarity 85.2%; Pred. No. 1.1e-93;  
Matches 453; Conservative 0; Mismatches 79; Indels 0; Gaps 0;  
QY 34 AAATTCATTAAGCAAGCAAGAAACACTTGTGTTAGTACATGCTGCATGCCATGGAGTT 93  
DB 4 AAGCTAACAGAAACCAAGGAAACAACTTCTGTTCTAGTACATGCTGCAGCCATGGAGCTT 63  
QY 94 GGAGTTGTGTACAGCTAAAGCCACCTGCTAGAGCTGCAGGCCATAAGGTTACAGCCCTTG 153  
DB 64 GGTGTTGTGTACAGCTAAACCAATGCTAGAGCTGCAGGCCACCAAGGTCACCTGCTTGT 123  
QY 154 ATTTAGCAGCTTCTGGCACTGATTTTGAAAAAATAGAGAGCTTCGCACTTTATGATT 213  
DB 124 ACTTAGCAGCTCTGGCATTGATTTTGAAAAAATAGAGCAACTTCACACACTTCATGATT 183  
QY 214 ATACTTTGCCATTGATGAGGTTGATGGAATCTCTTTACGACATGAGAGGTTATATTAG 273  
DB 184 ACACCTTTGCCATTATTTGGAATTTAATGGAATCTCTTCCACAAGAGGAGAAAGCCATTATTAG 243  
QY 274 TGGGCACTAGTCTTGGTGTATGAATTTGGCACTTGTCTATCGAAAAGTATCCACAAAAAGA 333  
DB 244 TTGGACATAGTCTTGGTGTATGAATTTAGCACTTGTCTATGGAAAAAATACCCAAAAAGA 303  
QY 334 TCTATGCTGCTGTTTTTCTTGGCTGCTTTTCACTGCCCTGATTCTGTTTCACTTCCTCTTG 393  
DB 304 TCTATGCTGCTGTTTTTCTTGGCTGCTTTTATGCTGCTGATTCTATTACACATATCTTCTCTATG 363  
QY 394 TTTTGGACAGTATAATGAGCGGACGCCAGCCAGAGAAATGCTTGGATCTACTCAGTTTTTAC 453  
DB 364 TTATGGATCAGTACAAATGAACGGACCAACGACAGAGAAATGGTTAGATACTCAATTTTTTAC 423  
QY 454 CATATGGTTCCTCCCTGAAGAGCCACTGACATCCATGTTTTTTTGGGCCCAAGTCTTCTGGGCTC 513

Db 424 CATATGGTACCCCTGAAGAGCCACTCATCATCCAGACTTTTGGTCCCAATTTTGGCTG 483

Qy 514 ACAAGCTCTACAGCATGCTCTCTCTGAGGATCTTGCATTAGCATCATCAT 565

Db 484 ATAAGCTTTACCGGTTAAGCCCTCTGGAAGATGTTGCATTACGATTATCAT 535

CV267710 957 bp mRNA linear EST 22-SEP-2004

WS02032\_B21\_H19 PTXN-IB-N-A-11 Populus balsamifera subsp. trichocarpa x Populus nigra cDNA clone WS02032\_H19 3', mRNA sequence.

CV267710

CV267710.1 GI:52520685

EST.

Populus balsamifera subsp. trichocarpa x Populus nigra

Populus balsamifera subsp. trichocarpa x Populus nigra

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Malpighiales; Salicaceae; Populus.

1 (bases 1 to 957)

Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y., Kirkpatrick,R., Liu,J., Palmquist,D., Scott,J., Barber,S., Yang,G., Babakalif,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A., Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A., Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and Bohlmann,J.

The poplar transcriptome: Analysis of expressed sequence tags from multiple cDNA libraries

Unpublished (2004)

Contact: Joerg Bohlmann

Genome BC forest genomics program

University of British Columbia

UBC Biotechnology Laboratory, 6174 University Boulevard, Rm. 237, Vancouver, British Columbia, Canada, V6T 1Z3

Tel: 1-604-822-0282

Fax: 1-604-822-6097

Email: bohlmann@interchange.ubc.ca

Plate: WS02032 row: H column: 19

High quality sequence stop: 957.

Location/Qualifiers

1..957

/organism="Populus balsamifera subsp. trichocarpa x Populus nigra"

/mol\_type="mRNA"

/cultivar="NxM6"

/db\_xref="taxon:293756"

/clone="WS02032\_H19"

/sex="Male"

/lab\_host="E. coli DH10B T1 phage resistant cells"

/clone\_lib="PTXN-IB-N-A-11"

/note="Vector: pBluescript II SK (+) XR; Site 1: EcoRI (5' end of cDNA); Site 2: XhoI (3' end of cDNA); Sapling trees two metres in height and grown under greenhouse conditions were exposed to continuous feeding by Cryptorhynchus lapathi (poplar and willow borer) adults caged on the sapling using mesh bags. Bark (with phloem and cambium attached) from within the caged region was harvested 2 hours, 6 hours and 48 hours after the onset of treatment. mRNA was isolated from each tissue source independently and equal quantities of mRNA from each tissue were then pooled. cDNA was prepared from 5 micrograms of mRNA and directionally ligated into the pBluescript II SK (+) XR vector using the pBluescript II XR cDNA Library Construction Kit according to manufacturer's instructions with modifications (Stratagene). Plasmid DNA was then transformed by electroporation into DH10B cells (Invitrogen) for propagation. Normalization was applied according to published methods [Bonaldo M.F. et al. (1996) Genome Research 6(9):791] in order to reduce the abundance of highly expressed transcripts."

| ORIGIN |     | Query Match   | 34.9%           | Score 376.4;       | DB 7;     | Length 957; |
|--------|-----|---|-----------------|--------------------|-----------|-------------|
|        |     | Best Local Similarity   | 67.1%           | Pred. No. 4.2e-86; |           |             |
|        |     | Matches 533;  | Conservative 0; | Mismatches 261;    | Indels 0; | Gaps 0;     |
| Qy     | 40  | AAATGAAGGAGGAAAAACACTTTGTTTGTAGTACATGTCATGTCATGCGATGGAGGTTTGGAGTT | 99              |                    |           |             |
| Db     | 154 | AGACCAAGATCAGNAGCATTTTGTCTAGTACATGAGCTTGCATGGGGCTTGTGTCT          | 213             |                    |           |             |
| Qy     | 100 | GGTCAAGCTAAAGCCCACTGCTAGAGCTGAGGCCATAAGGTTACAGCCCTTCAATTAG        | 159             |                    |           |             |
| Db     | 214 | GGCAAAAGTTCAAAAACGCTGCTTGAGTCAGCAAGTAACAGGGTCACGGTCTTGACCTTG      | 273             |                    |           |             |
| Qy     | 160 | CAGCTTTCGGCACTGATTTTGAGAAAAATAGAGGAGCTTCGCACACTTTTATGATTAATCTT    | 219             |                    |           |             |
| Db     | 274 | CTGCTTCAGGCGCCAAACATGAAGGCAATCCAAAGATGTAGAAACACTTGTATGAATATACGG   | 333             |                    |           |             |
| Qy     | 220 | TGCCATTGATGGAGTTGATGGAATCTCTTTTCAGCAGATAGAGGTTTATATTAGTGGGGC      | 279             |                    |           |             |
| Db     | 334 | AGCCTTTGTTAGAGTTTCTGCGCTCATTAACCGAGGAGAGGTCATTTCTAGTAGGGC         | 393             |                    |           |             |
| Qy     | 280 | ATAGTCTTGGTGGTATGAATTTGGGACTTGCTATGGAAGATATCCACAAAGATCTATG        | 339             |                    |           |             |
| Db     | 394 | ACAGCTTAGGAGGTTTGAGTTTGGCTCTTGCTATGGAAGTTCCCGAGAGAGATTTGCTG       | 453             |                    |           |             |
| Qy     | 340 | CTGCTGTTTCTTGGCTGCTTTTCATGCTGATTTCTGTTTTCACAACTCCCTCTTCTTTGG      | 399             |                    |           |             |
| Db     | 454 | TGCTGTTTCTTATCAGCTTTTCATGCGAGATACACACAAAGCCATCATTTGTCTTGG         | 513             |                    |           |             |
| Qy     | 400 | AACAGTATAATGAGCGGACCGCAGCGAGAAATTTGGTTGGATCTCAGTTTTTACCATATG      | 459             |                    |           |             |
| Db     | 514 | ATCAGTATACGAGAGAGACCCCGGGGATTCCTGTTGGACACTCAATTTTACCNTACA         | 573             |                    |           |             |
| Qy     | 460 | GTTCCCTCGAAGAGCCACTGACATCCATGTTTTTTTGGCCCAAGTTCTTGGCTCACAAGC      | 519             |                    |           |             |
| Db     | 574 | GCAGTTCTCAAAGTCATCTCACAACAATGTCTTTGGACCCAAATTTCTTATCTCTCCAAGC     | 633             |                    |           |             |
| Qy     | 520 | TCTACGACTATGCTCTCTCGAGGATCTTGCAATTAGCATCATCTGGTGAGACCAAGCT        | 579             |                    |           |             |
| Db     | 634 | TCTATCAGCTAAGCCCACTTGAGGATCTTGAGCAAGCAAGACTATGGTAAAGGCCAGAT       | 693             |                    |           |             |
| Qy     | 580 | CTTTGTTTATGGAAGACCTATCGAAGCCAAAGTATTTTCACAGATGAACGGTTTGGATCAG     | 639             |                    |           |             |
| Db     | 694 | CACGTTTCTGTATGATTTTGTCAAAGGCAACAGTTTCTCCACGCGGCTATGGGTCAG         | 753             |                    |           |             |
| Qy     | 640 | TGAAGAGAGTTTACATTGTGTGCACTGAGGATAAAGGCATACCGAAGAATTCAGCGAT        | 699             |                    |           |             |
| Db     | 754 | TCAACCGAGTGTATGTTTATCTGCGATGAAGATTTTAGCGATACCGAAGAGTTTCAACGCT     | 813             |                    |           |             |
| Qy     | 700 | GGCAAAATTGCAACATTTGGTCTCACTGAAGCAATAGAGATTAAAGGTCTGATCACATGG      | 759             |                    |           |             |
| Db     | 814 | GGATGATTGAAACACAGTCTGTGAAGAAAGTTATGGAATTTGAAGGTGCACCAATATGG       | 873             |                    |           |             |
| Qy     | 760 | CAATGCTATGCGAGCCCAAAACCTTTTCGCGCTCTCTCTTGGAAATTTGCCCAATAATACA     | 819             |                    |           |             |
| Db     | 874 | TTATGTTCAGAGGCCACCAAGACTCTTCCATTTGTCTCTCAGAGTACGAAATAACATG        | 933             |                    |           |             |
| Qy     | 820 | ACTGATCTCTACAT  | 833             |                    |           |             |
| Db     | 934 | CTTAAGCTGTACTT  | 947             |                    |           |             |

Search completed: September 12, 2005, 06:31:50  
Job time : 3448.23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 01:44:06 ; Search time 205.395 Seconds  
(without alignments)  
8595.849 Million cell updates/sec

Title: US-10-780-002-1  
Perfect score: 1079  
Sequence: 1 acgcggggaagaaagaaaa.....gaaaaaaaaaaaaaaaaa 1079

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/PTUS\_COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description          |
|------------|-------|-------------|--------|----|----------------------|
| 1          | 186   | 17.2        | 1078   | 3  | US-08-981-256A-4     |
| 2          | 49    | 4.5         | 7218   | 1  | US-08-232-463-14     |
| 3          | 44.8  | 4.2         | 1141   | 4  | US-09-806-708B-22    |
| C 4        | 40    | 3.7         | 601    | 4  | US-09-949-016-153422 |
| C 5        | 40    | 3.7         | 601    | 4  | US-09-949-016-153423 |
| 6          | 40    | 3.7         | 130563 | 4  | US-09-949-016-12273  |
| 7          | 40    | 3.7         | 131379 | 4  | US-09-949-016-16050  |
| C 8        | 39.2  | 3.6         | 83938  | 4  | US-09-949-016-16068  |
| 9          | 38.8  | 3.6         | 601    | 4  | US-09-949-016-154474 |
| 10         | 38.8  | 3.6         | 1782   | 4  | US-09-248-796A-369   |
| 11         | 38.6  | 3.6         | 2475   | 4  | US-09-220-132-179    |
| 12         | 38.6  | 3.6         | 2884   | 3  | US-09-439-313-335    |
| 13         | 38.6  | 3.6         | 2984   | 3  | US-09-352-616A-335   |
| 14         | 38.6  | 3.6         | 2984   | 3  | US-09-232-149A-335   |
| 15         | 38.6  | 3.6         | 2984   | 4  | US-09-636-215-335    |
| 16         | 38.6  | 3.6         | 2984   | 4  | US-09-685-166A-335   |
| 17         | 38.6  | 3.6         | 2984   | 4  | US-09-688-489-335    |
| 18         | 38.6  | 3.6         | 2984   | 4  | US-09-679-426-335    |
| 19         | 38.6  | 3.6         | 2984   | 4  | US-09-759-143-335    |
| 20         | 38.6  | 3.6         | 2984   | 4  | US-09-651-236-335    |
| 21         | 38.6  | 3.6         | 60593  | 4  | US-09-949-016-11779  |
| 22         | 38    | 3.5         | 32798  | 4  | US-09-949-016-11758  |
| 23         | 38    | 3.5         | 32798  | 4  | US-09-949-016-11758  |
| C 24       | 37.8  | 3.5         | 1141   | 4  | US-09-806-708B-22    |
| C 25       | 37.8  | 3.5         | 85850  | 4  | US-09-949-016-13424  |
| 26         | 37.6  | 3.5         | 2183   | 3  | US-09-352-990-13     |
| 27         | 37.6  | 3.5         | 42000  | 4  | US-10-081-563-25     |
| 28         | 37.6  | 3.5         | 392000 | 4  | US-10-027-983-11     |
| C 29       | 37.2  | 3.4         | 219964 | 4  | US-09-949-016-15086  |
| 30         | 37    | 3.4         | 832    | 4  | US-09-621-976-2813   |
| 31         | 37    | 3.4         | 1929   | 2  | US-08-892-770-2      |
| 32         | 36.6  | 3.4         | 1793   | 4  | US-09-620-312D-996   |
| C 33       | 36.6  | 3.4         | 7218   | 1  | US-08-232-463-14     |
| 34         | 36.6  | 3.4         | 23497  | 4  | US-09-949-016-117072 |
| 35         | 36.6  | 3.4         | 149971 | 4  | US-09-949-016-113590 |
| C 36       | 36.4  | 3.4         | 406    | 4  | US-09-829-481-3      |
| C 37       | 36.4  | 3.4         | 1362   | 4  | US-09-710-279-803    |
| C 38       | 36.4  | 3.4         | 1386   | 3  | US-09-134-001C-767   |
| 39         | 36.4  | 3.4         | 3553   | 4  | US-09-710-279-4250   |
| 40         | 36.4  | 3.4         | 4095   | 4  | US-09-710-279-4079   |
| C 41       | 36.2  | 3.4         | 132    | 4  | US-10-021-338A-66    |
| C 42       | 36.2  | 3.4         | 6350   | 2  | US-08-385-335A-9     |
| C 43       | 36    | 3.3         | 601    | 4  | US-09-949-016-32018  |
| C 44       | 36    | 3.3         | 601    | 4  | US-09-949-016-70476  |
| C 45       | 36    | 3.3         | 60002  | 4  | US-09-949-016-12464  |

RESULT 1

US-08-981-256A-4  
; Sequence 4, Application US/08981256A  
; Patent No. 6046042  
; GENERAL INFORMATION:

; APPLICANT: Meinhard HASSLACHER et al.  
; TITLE OF INVENTION: (S)-HYDROXY-NITRILE-LYASE FROM HEVEA BRASILIENSIS

; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.  
; STREET: 2033 K Street, N.W., Suite 800  
; CITY: Washington  
; STATE: D.C.

; COUNTRY: U.S.A.  
; ZIP: 20006

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS

; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/981-256A

; FILING DATE: December 22, 1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER:

; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew Jacob

; REGISTRATION NUMBER: 25,154  
; REFERENCE/DOCKET NUMBER: 1553-OZ1112  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 721-8200

; TELEFAX: (202) 721-8250  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1078 base pairs  
; TYPE: nucleic acid

; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-981-256A-4

Query Match 17.2%; Score 186; DB 3; Length 1078;  
Best Local Similarity 53.2%; Pred. No. 3.5e-44;

Matches 421; Conservative 0; Mismatches 365; Indels 6; Gaps 1;

QY 38 TCAGTAAAGGAAGAAACACTTTGTTTATGATCGATGCCATGGAGTTTGAG 97  
Db 37 TCAGAAATGGCATTCGCTCATTTTCTTTATTCATACCATATGCCACGGTGCATGGAT 96

```
QY 98 TTGGTCAAGCTAAAGCCACTGCTAGAAAGCTCAGGCCAATAGGTTACAGCCCTTGATT 157
Db 97 TTGGCAAGCTCAAAACCCCTCTTGAGGCACTTGCGCCACAAGGTTACTGCACTGCACT 156
QY 158 AGCAGCTTCGCACTGATTTGAGAAATAGAGGAGCTTCGCACATTTATGATTATAC 217
Db 157 TGCAGCAAGCGGCTTGACCCAAAGGCAATGAGGAGATTGGCTCATTTGATGATGATTC 216
QY 218 TTTCGATTCGATGAGGATTCATGGAATCTCTTTACGACATGAGAAGGTTATATTAGTGG 277
Db 217 TGAACCTTGTGACGTTCTTGAGGCACTCCCTCCAGGGGAAAGGTCATTCGTGGTGG 276
QY 278 GCATAGCTTTGGTGTATGAATTTGGGACTTGTCTATGGAAGGATCCACAAAAGATCTA 337
Db 277 CGAGAGCTGCGAGGACTCAATATAGCAATTCGCTGATATAATCTGTGAAAGATTGC 336
QY 338 TCTGCTGTTTCTTGGCTGCTTTTCATGCTGATTCGTGTCACAACTCTCTCTTTGTTT 397
Db 337 AGCTGCTGTTTCCCAATTCAGTATTTGCCAGACACCGAGCACTGCCCATCTTACGTCGT 396
QY 398 GGAAACAGTATATAGCGGACGCCAGCCGAGCAATGTTGGATATCTCACTTTTATACCATA 457
Db 397 GATAGCTCATGAGGCTGTTTCCCG-----ACTGGAAGACACACGATTTTACGTA 450
QY 458 TGGTTCCCTGAAGACCCACTGACATCCATGTTTGTGGCCCAAGTTCCTGGCTCAAA 517
Db 451 CACTAAAGATGCAAGGAGATAACTGGATTGAAACTGGGCTTCAGGCTTCGAGGGAATA 510
QY 518 GCTCTACAGCTATGCTCTCTGAGGATCTTCGATAGCATCATCATTTGGTGAGCCAAAG 577
Db 511 TTATATACCTTTGCGGCTCGTGAAGATATGAACCTGGCGAAGATGTTGACAAGGAAGG 570
QY 578 CTCTTTGTTTATGGAAGACTATCGAAGGCCAAGTATTTACAGATGAACGGTTTGGATC 637
Db 571 ATCATTTTCAAATATTTTACCTTAGCGACCATCTTCACTAAGGAAGTTAGGATC 630
QY 638 AGTGAAGAGAGTTTACATTTGTGCACTGAGGATAAAGGCATACAGAGAATTCGAGG 697
Db 631 GATTAAGAAATTTATGTGTGACCGACCAAGACGAAATATTTTACCTGAAATTTCACT 690
QY 698 ATGGCAATGACAACTTGTGTCTACTGAAGCAATAGAGATTAAGGTGCTGATCATAT 757
Db 691 CTGGCAATAGAAAATCTATAACCAAGCAAGGTTTATAAGGTGCGAAGGTGAGATCATAA 750
QY 758 GGCAATGCTATGCGACCCCAAAAACCTTTGCGCTCTCTTTGGAAATTCGCCATAAATA 817
Db 751 ATTGCACTTACAAAGACTAAGGAGATCGCTGAAATCTCCAAAGAGTGCTGATACCTA 810
QY 818 CAACGATCTCT 829
Db 811 TAATTGACTTCT 822
```

## RESULT 2

```
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
```

## GENERAL INFORMATION:

```
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
US-08-232-463-14
```

Query Match 4.5%; Score 49; DB 1; Length 7218;

Best Local Similarity 7.4%; Pred.No.0.0016;

Matches 25; Conservative 177; Mismatches 137; Indels 0; Gaps 0;

```
QY 721 TCACATGAAGCAATAGAGATTAAAGTGCTGATCATGCGCAATGCTATGCGAGCCCAAA 780
Db 1019 TCACAGATTAAATTCGAGCTTGCTGAGCTCGAGGAGCTTGCAATGATGATGATGAT 1078
QY 781 AACTTTGCGCTCTCTCTTGGAAATGCGCAATAAATCAACTGATCTCTCATGATGCT 840
Db 1079 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1138
QY 841 TCGTCTCATGTCAGATTTTCAGTGCATGCTGTAATTTTCTATTTTTCGACGCGCGC 900
Db 1139 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1198
QY 901 ATAACGTGCTTGCCTATTAAAGGATTGCAAGTAAATTCACCTCTCTAGTGTGAAGCT 960
Db 1199 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1258
QY 961 TCCACATAAGAGATTGCTGTTCTCCATTCCCAATCAAGTGTGTTATGTTGAGATACCTA 1020
Db 1259 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1318
QY 1021 CGTATCAATTCGTAAATGAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1059
Db 1319 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1357
```

## RESULT 3

US-09-806-708B-22

```
; Sequence 22, Application US/09806708B
; Patent No. 6784342
```

## GENERAL INFORMATION:

```
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
```





```
RESULT 9
US-09-949-016-154474
; Sequence 154474, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 154474
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-154474

Query Match      3.6%; Score 38.8; DB 4; Length 601;
Best Local Similarity 46.8%; Pred. No. 0.4; Indels 0; Gaps 0;
Matches 118; Conservative 1; Mismatches 133;

QY 822 TGATCTCTACATATATGTCCTCGTCTCATGTCAAGATTTTCAGTGCATGCTGTAATTTTT 881
DB 148 TTACTTCTACAGTTGTTCTTGTCATTAATGCTGTAATTTATAAAGTGTCAAGTCTTG 207

QY 882 TCTATTTTTCGACCGCGGCACTACTGCTTTGCTCTATTTTAAGGATTCGAGTAATTTTCA 941
DB 208 ACCTGCTGTGTGACCACTAATATGTACCTAGCTTTGGCTAAGTCTTTAACTAATTTATAC 267

QY 942 TCTTCTAGTGTGAAGGCTTCCACATAAGGATTTGTTCTGTTCTCCATTCAGTGTGCT 1001
DB 268 TCTAAATGATATAGCTAATATTTACTGAGGCTTGCTATATGCGAGTACTTTCTGA 327

QY 1002 TATGTTGAGATACTTAAACCGTATCAATTTCTGTATGAACACTTCTTCTTCTTTTGA 1061
DB 328 GTGTTTCACATATTTAGACCCATTTATCTCCATGAGTAGGCATTAATTTCTATTTA 387

QY 1062 AAAAAAAAAA 1073
DB 388 TAGATAAAGAAA 399

RESULT 10
US-09-248-796A-369
; Sequence 369, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 369
; LENGTH: 1782
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-369
```

```
Query Match      3.6%; Score 38.8; DB 4; Length 1782;
Best Local Similarity 64.4%; Pred. No. 0.73;
Matches 58; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 645 AGAGTTTACATTTGTCGCACTGAGGATTAAGGCATACCAAGAAGATTCAGCGATGGCAA 704
DB 1219 AAGATATCAATCTTGGACACTGACAAATAAAGCACATCCACGAATTTAAATAGTCAGAAA 1278

QY 705 ATTGACAACATTTGGTGTCACTGAACAATA 734
DB 1279 ATTAATACAGTGGTGATATGATTTCAATA 1308

RESULT 11
US-09-220-132-179
; Sequence 179, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 179
; LENGTH: 2475
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-179
```

```
Query Match      3.6%; Score 38.6; DB 4; Length 2475;
Best Local Similarity 45.5%; Pred. No. 1;
Matches 137; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 775 CCCAAAACTTGGCGCTCTCTCTGGAAATGGCCATAAATACAACTGATCTCTACATT 834
DB 2074 CACTTAAATTTGGAACCTTAAAGTTTAGATGCAATGTATATAAAAAAACCTTAAAGCAGTATC 2133

QY 835 ATGCTCTGCTCATGTCAGAGATTTTCAGTGCATGCTGTAATTTTTTCTATTTTCGAC 894
DB 2134 TGTATTTAGCTGTAAACCAAGTGGAAAGCTATTCCGATAATTTCTTAAATATGATGAA 2193

QY 895 CGGCGCATAAAGTCTTTTGGCCTATTTTAAGGATTCAGTAATTTTCACTCTTCTAGTGG 954
DB 2194 CTTTGGAGTACTGTTTCTTCTTCAAACTGAATGTAATTAATTCATGAATAATGCACCT 2253

QY 955 AAGCTTCCACATAAGGATTTGTTTCTTCCATTCAGTGTGTTGTTATGTTGATAC 1014
DB 2254 TATATGTTTAAACAATCTTTGTATACATTTTGGGATTTTGGTGTCTATATGCTAAATCAC 2313

QY 1015 TTAAACCTATCAATCTTGTGAAGAACTTCTTCTTCTTCTTTTGAAGAAAAA 1074
DB 2314 ATTCAGCATGTATTTTGACATTTTAAATACTTCCCTCAATTCGTAAATTTAAAGAAT 2373

QY 1075 A 1075
DB 2374 A 2374
```

```
RESULT 12
US-09-439-313-335
; Sequence 335, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
```



```
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqi
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 335
; LENGTH: 2984
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-439-313-335

Query Match          3.6%; Score 38.6; DB 3; Length 2984;
Best Local Similarity 56.8%; Pred. No. 1.1;
Matches 71; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 955 AAGGCTTCACATAAGGATGTTCTGTTCTCCATTCAGTGTGTTATGTTGAGATAC 1014
Db 2855 AAGTCTGTGACTGGAAGTTTAGTCAGCACCCGCCAACACTTATTTTCTATGTGTTT 2914

QY 1015 TTAACCGTATCAATCTTGTATGAACCTCTCTTTCTTTTGAAGAAAAA 1074
Db 2915 TTGCAACATATGAGTGTGTTTGAATAAAGTACCCATGCTTTTATTAGAAAAA 2974

QY 1075 AAAAA 1079
Db 2975 AAAAA 2979

RESULT 13
US-09-352-616A-335
; Sequence 335, Application US/09352616A
; Patent No. 6395278
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang, Yuqi
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF PROSTATE CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427C8
; CURRENT APPLICATION NUMBER: US/09/352,616A
; CURRENT FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 335
; LENGTH: 2984
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-352-616A-335

Query Match          3.6%; Score 38.6; DB 3; Length 2984;
Best Local Similarity 56.8%; Pred. No. 1.1;
Matches 71; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 955 AAGGCTTCACATAAGGATGTTCTGTTCTCCATTCAGTGTGTTATGTTGAGATAC 1014
Db 2855 AAGTCTGTGACTGGAAGTTTAGTCAGCACCCGCCAACACTTATTTTCTATGTGTTT 2914

QY 1015 TTAACCGTATCAATCTTGTATGAACCTCTCTTTCTTTTGAAGAAAAA 1074
Db 2915 TTGCAACATATGAGTGTGTTTGAATAAAGTACCCATGCTTTTATTAGAAAAA 2974

QY 1075 AAAAA 1079
Db 2975 AAAAA 2979

RESULT 14
US-09-232-149A-335
; Sequence 335, Application US/09232149A
; Patent No. 6465611
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE
; TITLE OF INVENTION: CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427C6
; CURRENT APPLICATION NUMBER: US/09/232,149A
; CURRENT FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 338
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 335
; LENGTH: 2984
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-232-149A-335

Query Match          3.6%; Score 38.6; DB 3; Length 2984;
Best Local Similarity 56.8%; Pred. No. 1.1;
Matches 71; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 955 AAGGCTTCACATAAGGATGTTCTGTTCTCCATTCAGTGTGTTATGTTGAGATAC 1014
Db 2855 AAGTCTGTGACTGGAAGTTTAGTCAGCACCCGCCAACACTTATTTTCTATGTGTTT 2914

QY 1015 TTAACCGTATCAATCTTGTATGAACCTCTCTTTCTTTTGAAGAAAAA 1074
Db 2915 TTGCAACATATGAGTGTGTTTGAATAAAGTACCCATGCTTTTATTAGAAAAA 2974

QY 1075 AAAAA 1079
Db 2975 AAAAA 2979

RESULT 15
US-09-636-215-335
; Sequence 335, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 335
```



```
; LENGTH: 2984
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-636-215-335

Query Match      3.6%; Score 38.6; DB 4; Length 2984;
Best Local Similarity 56.8%; Pred. No. 1.1;
Matches 71; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY  955 AAGGCTTCCACATAAGGATTGTTCTGTTTCTCCATTCAAGTGTGTTATGTGAGATAC 1014
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  2855 AAGTCTGTGACTTGAAGTTTAGTCAGCACCCCCACCACCAACTTTATTTTCTATGTGTTT 2914
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  1015 TTAACCGTATCAATTCTTGTAAATGAAACTTCTTCTTCTTTTGAAAAA 1074
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  2915 TTGCAACATATGAGTGTGTTTGAAATAAAGTACCCTGCTTTATTAGAAAAA 2974
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  1075 AAAAA 1079
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  2975 AAAAA 2979
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: September 12, 2005, 06:38:17  
Job time : 208.395 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 04:46:01 ; Search time 2389.95 Seconds  
(without alignments)  
2964.938 Million cell updates/sec

Title: US-10-780-002-1  
Perfect score: 1079  
Sequence: 1 acgcgggaagaaagaaaa.....gaaaaaagaaaaa 1079

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7351250 seqs, 3283620254 residues

Total number of hits satisfying chosen parameters: 14702500

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:\*
- 22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq:\*
- 24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description         |
|------------|-------|-------------|--------|----|---------------------|
| 1          | 1079  | 100.0       | 1079   | 21 | US-10-780-002-1     |
| 2          | 320   | 29.7        | 933    | 18 | US-10-424-599-64748 |
| 3          | 295.2 | 27.4        | 989    | 21 | US-10-780-002-32    |
| 4          | 292.2 | 27.1        | 792    | 9  | US-09-938-842A-1320 |
| 5          | 292.2 | 27.1        | 792    | 11 | US-09-938-842A-1320 |
| 6          | 277.6 | 25.7        | 1158   | 18 | US-10-424-599-85023 |
| 7          | 274.6 | 25.4        | 870    | 21 | US-10-780-002-33    |

|    |       |      |      |    |                      |                    |
|----|-------|------|------|----|----------------------|--------------------|
| 8  | 271.6 | 25.2 | 837  | 21 | US-10-780-002-34     | Sequence 34, Appl  |
| 9  | 271   | 25.1 | 792  | 9  | US-09-938-842A-1551  | Sequence 1551, Ap  |
| 10 | 271   | 25.1 | 792  | 11 | US-09-938-842A-1551  | Sequence 1551, Ap  |
| 11 | 269   | 24.9 | 792  | 9  | US-09-938-842A-217   | Sequence 217, App  |
| 12 | 269   | 24.9 | 792  | 11 | US-09-938-842A-217   | Sequence 217, App  |
| 13 | 268.2 | 24.9 | 1035 | 18 | US-10-425-114-12800  | Sequence 12800, A  |
| 14 | 256   | 23.7 | 792  | 21 | US-10-780-002-35     | Sequence 35, Appl  |
| 15 | 252.6 | 23.4 | 819  | 21 | US-10-780-002-37     | Sequence 37, Appl  |
| 16 | 240.8 | 22.3 | 653  | 18 | US-10-424-599-52897  | Sequence 52897, A  |
| 17 | 240.8 | 22.3 | 834  | 21 | US-10-780-002-39     | Sequence 39, Appl  |
| 18 | 239.8 | 22.2 | 771  | 9  | US-09-938-842A-1278  | Sequence 1278, Ap  |
| 19 | 239.8 | 22.2 | 771  | 11 | US-09-938-842A-1278  | Sequence 1278, Ap  |
| 20 | 239.8 | 22.2 | 771  | 21 | US-10-780-002-40     | Sequence 40, Appl  |
| 21 | 232.6 | 21.6 | 912  | 21 | US-10-780-002-36     | Sequence 36, Appl  |
| 22 | 224.8 | 20.8 | 783  | 21 | US-10-780-002-38     | Sequence 38, Appl  |
| 23 | 215   | 19.9 | 901  | 21 | US-10-780-002-41     | Sequence 41, Appl  |
| 24 | 197.8 | 18.3 | 920  | 18 | US-10-424-599-128189 | Sequence 128189, A |
| 25 | 195.6 | 18.1 | 1101 | 19 | US-10-437-963-62048  | Sequence 62048, A  |
| 26 | 191.8 | 17.8 | 901  | 20 | US-10-739-930-3765   | Sequence 3765, Ap  |
| 27 | 178.6 | 16.6 | 439  | 18 | US-10-424-599-81721  | Sequence 81721, A  |
| 28 | 175.2 | 16.2 | 1069 | 19 | US-10-767-701-9317   | Sequence 9317, Ap  |
| 29 | 169   | 15.7 | 777  | 21 | US-10-738-922-1      | Sequence 1, Appli  |
| 30 | 169   | 15.7 | 777  | 21 | US-10-738-927-1      | Sequence 1, Appli  |
| 31 | 156.6 | 14.5 | 777  | 21 | US-10-738-922-2      | Sequence 2, Appli  |
| 32 | 156.6 | 14.5 | 777  | 21 | US-10-738-927-2      | Sequence 2, Appli  |
| 33 | 154.6 | 14.3 | 1007 | 18 | US-10-425-114-21611  | Sequence 21611, A  |
| 34 | 151.6 | 14.1 | 1018 | 20 | US-10-425-114-119846 | Sequence 119846, A |
| 35 | 146   | 13.5 | 1025 | 18 | US-10-425-114-13774  | Sequence 13774, A  |
| 36 | 146   | 13.5 | 1092 | 20 | US-10-425-115-70617  | Sequence 70617, A  |
| 37 | 145.6 | 13.5 | 1165 | 18 | US-10-424-599-62644  | Sequence 62644, A  |
| 38 | 133   | 12.3 | 1039 | 18 | US-10-425-114-24225  | Sequence 24225, A  |
| 39 | 133   | 12.3 | 1249 | 20 | US-10-425-115-179879 | Sequence 179879, A |
| 40 | 131.8 | 12.2 | 1148 | 19 | US-10-437-963-29636  | Sequence 29636, A  |
| 41 | 128.2 | 11.9 | 996  | 19 | US-10-767-701-15339  | Sequence 15339, A  |
| 42 | 126.2 | 11.7 | 483  | 10 | US-09-770-961-854    | Sequence 854, App  |
| 43 | 125.2 | 11.6 | 1155 | 15 | US-10-163-198-100    | Sequence 100, App  |
| 44 | 120.6 | 11.2 | 1584 | 18 | US-10-424-599-82031  | Sequence 82031, A  |
| 45 | 110.6 | 10.3 | 1006 | 20 | US-10-425-115-36145  | Sequence 36145, A  |

ALIGNMENTS

RESULT 1  
US-10-780-002-1  
; Sequence 1, Application US/10780002  
; Publication No. US20050034196A1  
; GENERAL INFORMATION:  
; APPLICANT: Klesig, Daniel F.  
; APPLICANT: Kumar, Dhirenda  
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN  
; TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, SABB2, AND METHODS OF USE THEREOF  
; FILE REFERENCE: 3670-FO2652WO  
; CURRENT APPLICATION NUMBER: US/10/780,002  
; PRIOR FILING DATE: 2004-02-17  
; PRIOR APPLICATION NUMBER: PCT/US02/26312  
; PRIOR FILING DATE: 2002-08-16  
; PRIOR APPLICATION NUMBER: 60/312,863  
; PRIOR FILING DATE: 2001-08-16  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Nicotiana tabacum  
US-10-780-002-1

Query Match 100.0%; Score 1079; DB 21; Length 1079;  
Best Local Similarity 100.0%; Pred. No. 4.1e-295;  
Matches 1079; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACGCCGGGAAAGAAAGAACTTAAACAGGCATTAATCAATGAAGGAAACACT 60  
|||||

1 ACGCGGGAAGAAAGAACTAAACAGGCATAAAATTCAAATGAAGGAAGAAACACT 60  
61 TTGTTTTAGTACATGGTGATGCCATGGAGGTTGGAGTTGGTACAAAGCTAAAGCCACTGC 120  
61 TTGTTTTAGTACATGGTGATGCCATGGAGGTTGGAGTTGGTACAAAGCTAAAGCCACTGC 120  
121 TAGAAGCTGCAGGCCATAAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCACTGATTTGA 180  
121 TAGAAGCTGCAGGCCATAAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCACTGATTTGA 180  
181 GAAATAATAGAGAGCTTGCACACTTTATGATTTATACCTTTGGCACTGATGAGTTGATGG 240  
181 GAAATAATAGAGAGCTTGCACACTTTATGATTTATACCTTTGGCACTGATGAGTTGATGG 240  
241 AATCTCTTTTACAGCATGAGAAGGTTATATTAGTGGGGCATAGTCTTGTTGGTATGAATT 300  
241 AATCTCTTTTACAGCATGAGAAGGTTATATTAGTGGGGCATAGTCTTGTTGGTATGAATT 300  
301 TGGGACTTGTCTATGGAAGATATCCAAAGATCTATGCTGCTGTTCTTGGCTGCTT 360  
301 TGGGACTTGTCTATGGAAGATATCCAAAGATCTATGCTGCTGTTCTTGGCTGCTT 360  
361 TCATGCTGATTTCTGTTTCAACTCTCTCTTTGTTGGCAAGATATATGAGCGGACGC 420  
361 TCATGCTGATTTCTGTTTCAACTCTCTCTTTGTTGGCAAGATATATGAGCGGACGC 420  
421 CAGCGAGAAATTTGGTTGGATCTCAAGTTTTCATATGTTTCCCTGAAAGCACTGA 480  
421 CAGCGAGAAATTTGGTTGGATCTCAAGTTTTCATATGTTTCCCTGAAAGCACTGA 480  
481 CATCCATGTTTTTGGCCCAAGATTTCTGGCTCACAGCTTACAGCTATGCTCTCCTG 540  
481 CATCCATGTTTTTGGCCCAAGATTTCTGGCTCACAGCTTACAGCTATGCTCTCCTG 540  
541 AGGATCTTCATTTAGCATCATGTTGGTGAGACCAAGCTCTTTGTTATGGAAGCACTAT 600  
541 AGGATCTTCATTTAGCATCATGTTGGTGAGACCAAGCTCTTTGTTATGGAAGCACTAT 600  
601 CGAAGGCCAAGTATTTACAGATGAACGGTTTGGTTCAGTGAAGAGGTTTACATTTGT 660  
601 CGAAGGCCAAGTATTTACAGATGAACGGTTTGGTTCAGTGAAGAGGTTTACATTTGT 660  
661 GCATGAGATTAAGCATACCAAGAAATTCAGCGATGCGCAATTTGACACATTTGGT 720  
661 GCATGAGATTAAGCATACCAAGAAATTCAGCGATGCGCAATTTGACACATTTGGT 720  
721 TCACTGAAGCAATAGAGATTAAGGTTGATGATCACAATGGCAATGCTATGCCAGCCCAA 780  
721 TCACTGAAGCAATAGAGATTAAGGTTGATGATCACAATGGCAATGCTATGCCAGCCCAA 780  
781 AACTTTGCGCCTCTCTCTTTGGAATTTGCCATAAATFACAACATGATCTTACATTTATGCT 840  
781 AACTTTGCGCCTCTCTCTTTGGAATTTGCCATAAATFACAACATGATCTTACATTTATGCT 840  
841 TGTCTCATGTCAAGATTTTTCAGTGATGCTGTAATTTTTTTTCTATTTTTCGACCGGCG 900  
841 TGTCTCATGTCAAGATTTTTCAGTGATGCTGTAATTTTTTTTCTATTTTTCGACCGGCG 900  
901 ATAACTGTCTTGGCTTATTTAAGGATTCAGTAAATTTTCACTCTCTAGTGTGGAAGGCT 960  
901 ATAACTGTCTTGGCTTATTTAAGGATTCAGTAAATTTTCACTCTCTAGTGTGGAAGGCT 960  
961 TCCATAAAGGATTTGTTCTGTTTCTCCATTCAGTGTGTTGTTATGTTGATCTTAAAC 1020  
961 TCCATAAAGGATTTGTTCTGTTTCTCCATTCAGTGTGTTGTTATGTTGATCTTAAAC 1020  
1021 CGTATCAATTTCTGTAATGAACCTTCTCTTCTCTTTTGAAGGATTAAGGTTGCTGATCA 1079  
1021 CGTATCAATTTCTGTAATGAACCTTCTCTTCTCTTTTGAAGGATTAAGGTTGCTGATCA 1079

; Sequence 64748, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 64748  
; LENGTH: 933  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_29480C.1  
US-10-424-599-64748

Query Match 29.7%; Score 320; DB 18; Length 933;  
Best Local Similarity 63.9%; Pred. NO. 9.1e-80;  
Matches 507; Conservative 0; Mismatches 275; Indels 12; Gaps 1;

QY 35 AATTCAAATGAAGNAGGAAACACACTTTGTTTGTAGTACATGGTGCATGCCATGGAGGTTG 94  
DB 24 AAAAAATAGTATAGATAGNAGCACCTTTGTTCTGGTGCATGGGCATGCCATGGAGCATG 83  
QY 95 GAGTTGGTACAAAGCTTAAAGCCACTGCTAGAAGCTGCAGGCCATAAGGTTACAGCCCTTGA 154  
DB 84 GTGTTGGTATGAAGCTCAAGCCACGCTTGGATCTGCTGGGCACAAGGTCACAGTCTTGA 143  
QY 155 TTTAGCAGCTTCTGGCACTGATTTGAGAAAAATAGAGAGCTTGCACACATTTATGATTA 214  
DB 144 CTTTGCAGCTTCTGGAGCCAAACATGAAGAAAAATGAAGATGTGACACTTTTTCACAGTA 203  
QY 215 TACTTTCGATTTGATGGAGTTGATGGATCTTTTACAGCATGAGAGGTTATATTAGT 274  
DB 204 TACTGAGCCTTTGTTGTTTCTATTGGACACAAATTCCTTCAATGAAAAGGTAGTCTAGT 263  
QY 275 GGGGCATAGTCTTGGTGTATGAATTTGGGACTTGTCTATGGAAGATATCCACAAAAGAT 334  
DB 264 TGGTCAAGCTTTGGAGGGCTGACATAGCACTTGCCTGAGAGAAATTTCCAGAAAAGT 323  
QY 335 CTATGCTGCTGTTTCTTGGCTGCTTTCATGCTGATTTCTGTTCAACTCTCTCTTTGT 394  
DB 324 AGCAGTTGGTGTGTTTCTTAAACAGCTTTTGTCTCCAGACGTTGAACACCACCTATATGT 383  
QY 395 TTTGGAACAGTATATAGCGGACGCCAGCGAGAAATTTGGTGTGATCTCAGTTTTTACC 454  
DB 384 CTTGGAAGATACAGCGAGAGACCCGTTAGCTGCTAGTGGTTAGACACTGAAATTTGCTCC 443  
QY 455 ATATGGTTCCCTGAAGAGCCACTGACATCATGTTTGGCCCAAGTCTTGGCTCA 514  
DB 444 AAGTGAAC-----AAAAACAATGTTCTTTGGCCCAACTCTTGTCCGA 491  
QY 515 CAAGCTCTACAGCTATGCTCTCTGAGGATCTTGCATTTAGCATCATCATTTGGTGAGACC 574  
DB 492 CAAGCTCTACCAACTATCCCAATTTGAGGATTTTGAATTTGGCCAAAGACTTTAATAAGGCC 551  
QY 575 AAGCTCTTGTGTTTGAAGAGCCCTATCGAAGGCCAAGTATTTTACAGATGAACGGTTGG 634  
DB 552 ATCATCACTCTTCATGGAAGACTTGACTAAACAAGAACTTCTCCAAGAGGATATGG 611  
QY 635 ATCAGTGAAGAGAGTTTACATTTGTGTGCACTGAGGATAAAGGCATACACAGAAATTCOA 694  
DB 612 GTCACTGCCAGCTGCTTATTTGTTGCACTGAGGACCTTGCAATTTCCATTGGAATATCA 671  
QY 695 GCGATGGCAAAATTGACAACTTTGGTGTCTCACTGAAGCAATAGAGATTAAGGTTGCTGATCA 754  
DB 672 GCTCTTCATGATCCAAAAATGTTGGGTTCAATGAAGTTGTAGAGATCAAAAGGCCAGACCA 731



[illegible]

RESULT 5  
US-09-938-842A-1320  
; Sequence 1320, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harder, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRI1300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1320  
; LENGTH: 792  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1320

|    | Query Match               | 27.1%   | Score 292.2;       | DB 11;    | Length 792; |
|----|---------------------------|---|--------------------|-----------|-------------|
|    | Best Local Similarity     | 63.1%;  | Pred. No. 6.4e-72; |           |             |
|    | Matches 485; Conservative | 0;  | Mismatches 278;    | Indels 6; | Gaps 2;     |
| Qy | 53                        | AAAACACTTTGTTTTAGTACATGGTGCGATGCCATGGAGGTGGTAGTTGGTACAAGCTAAA   | 112                |           |             |
|    |                           |   |                    |           |             |
| Db | 21                        | ACAACATTTCGTACTAGTACATGGTTCGTGCCATGGCGGTGGTGGTACAAGGTTAA        | 80                 |           |             |
|    |                           |   |                    |           |             |
| Qy | 113                       | GCCACTGCTAGAAGCTGCAGGCCCATAAAGTTTACAGCCCTTCATTTAGCAGCTTCTGGCAC  | 172                |           |             |
|    |                           |   |                    |           |             |
| Db | 81                        | GCCGCTGCTAGAGGCGGTGGSCCACC GCGTAAC TGTCTGGACTTAGCTGCCTCCGGAAT   | 140                |           |             |
|    |                           |   |                    |           |             |
| Qy | 173                       | TGA---TTTGAGAAAAATAGAGGAGCTTFCGCACCTTTATGATTATAC TTTTGGCAATTGAT | 229                |           |             |

|    |     |          |           |          |           |          |         |           |           |                   |     |
|----|-----|----------|-----------|----------|-----------|----------|---------|-----------|-----------|-------------------|-----|
| Db | 141 | AGACACAA | CGAGGTG   | CGATCACT | CACTGACAT | CCCCACAT | CGGACAA | CAATACT   | CGGAGCC   | ATTGAC            | 200 |
| Qy | 230 | GGAGTTG  | ATGGAAT   | CTCTTT   | CACGAGAT  | GAGAAGT  | TATAT   | TAT       | TAGTGGG   | CATAGTCTTGG       | 289 |
| Db | 201 | GAAGCTC  | CTGACCT   | CATT     | TGCCAAA   | TGATG    | CAAAAGG | TTG       | TGCTCGT   | TGGTTCACAGCTTTGG  | 260 |
| Qy | 290 | TGTTATG  | AAATTTGGG | ACTTGCT  | TATG      | GGAAAAGT | TATCC   | CAAAAAGAT | CTATGCT   | GTCTTTT           | 349 |
| Db | 261 | TGGCTTG  | AACTTAG   | CCATAG   | CCATAG    | CCATAG   | CCATAG  | CCATAG    | CCATAG    | CCATAG            | 320 |
| Qy | 350 | CTTGGCT  | GTCTTT    | CATGCT   | CTGATTT   | CTGTT    | CAAACT  | CTCTCT    | TTGTTT    | TGGAAACAGTATAA    | 409 |
| Db | 321 | CTTGACT  | GTCTTT    | CTATG    | CCGGA     | CACCG    | AACTC   | CACTC     | CACTC     | CACTC             | 380 |
| Qy | 410 | TGACGGG  | AGCCG     | CCAGC    | CGAGAA    | TTGGT    | TGGAT   | ACTCAG    | ATTTTT    | TACCATATGGTTC     | 469 |
| Db | 381 | AAGCAAT  | GTGCTT    | CAAGA    | AGCAT     | TGGAT    | GGGCA   | CCGAA     | TTCGAA    | CCCTTATGGTT       | 440 |
| Qy | 470 | AGAGC    | CACTG     | CATCAT   | GTTTTT    | TGGCC    | CAAAAGT | TTCTT     | GGCTC     | ACAAGCTCTACAGCT   | 529 |
| Db | 441 | TTCCGGA  | CTGA      | ---GT    | ATGTTTT   | TTAG     | CCCTG   | CACTT     | CAATG     | AAAGTTGGGTCTCTA   | 497 |
| Qy | 530 | ATGCTCT  | CTCTG     | AGGAT    | CTTTG     | CATTAG   | CATCAT  | CATTGGT   | GAGACCA   | AAAGCTCTTTGTTTAT  | 589 |
| Db | 498 | TTCTCAG  | TTGAGAT   | CTTTG    | AATCGG    | ATTA     | CTTTT   | TAATG     | AGCCAG    | ATCGTTATTTAT      | 557 |
| Qy | 590 | GGAAG    | ACTAT     | CGAAG    | CCAAAGT   | ATTTT    | CAAGAT  | GAA       | CGGTTT    | TGATCAAGTGAAGAGAT | 649 |
| Db | 558 | TAACGAT  | TTTAT     | CGAAG    | TGAAAA    | CTTCT    | CGGAT   | GAA       | GGATAT    | GGGTCTGTTCCTCGAGT | 617 |
| Qy | 650 | TTACAT   | TGTGTG    | CACTG    | AGGAT     | TACCG    | AGGAA   | TTTCC     | AGCGAT    | TGGCAAAATGA       | 709 |
| Db | 618 | TTTCTAG  | TGTGT     | AAAG     | GAGCA     | AAAGCA   | ATTC    | CAAGA     | AAACG     | CCAGAGATGGATGATGA | 677 |
| Qy | 710 | CAACAT   | TTGGT     | GTCT     | CACTG     | AAAGCA   | TAGAGAT | TAAAG     | GTGCTG    | ATCATGGCAATGCTATG | 769 |
| Db | 678 | TAATTTT  | CCGGT     | GAAAT    | TTTAGT    | GNATG    | AGATG   | GAGG      | GAGACAGAT | CATATGCCAATGTTCTG | 737 |
| Qy | 770 | CGAGCCC  | CAAAA     | ACTTTG   | CGCCCT    | CTCTCT   | TTG     | GGAAAT    | TGCCC     | ATAAATAC          | 818 |
| Db | 738 | CAAGCCT  | CAGCA     | ACTCAGT  | GATTA     | CTTCT    | CGT     | GAATA     | TCCG      | GACAAATTC         | 786 |

RESULT 6

US-10-424-599-85023

; Sequence 85023, Application US/10424599

; Publication No. US20040031072A1

GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 85023

; LENGTH: 1158

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_47790C.1

US-10-424-599-85023

RESULT 6  
 US-10-424-599-85023  
 ; Sequence 85023, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 85023  
 ; LENGTH: 1158  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_47790C.1  
 US-10-424-599-85023

```

Query Match      25.7%; Score 277.6; DB 18; Length 1158;
Best Local Similarity 60.8%; Pred. No. 1.1e-67;
Matches 495; Conservative 0; Mismatches 304; Indels 15; Gaps 2;

Qy 1  ACGCGGGGAAGAAAGAAACTAACAGGCATTAATTTCAATTAAGACGAGGAAACACT 60
    152  ACGGAGCAAAAGCGTAGAAAGCGGCGAGTATGATTAACAAACAAAGAGAGCA 208
    AATCACT 208

```

QY 61 TTGTTTTAGTACATGGTGCATGCGAGGTTGGAGTTGGTACAGCTAAAGCCACTGC 120  
Db 209 TTGCTCTGGTGCATGGTATAGGCATGGTCCCTGGTGTGGTACAGCTTAAAGCCACTGT 268  
QY 121 TAGAAGCTGCAGGCCATAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCAGCTGATTTGA 180  
Db 269 TGGAAATCCGCGGCCACAAAGTCACAGCTTGCAGCTTCTGGCATCGACACAC 328  
QY 181 GAAATATAGAGGAGCTTCGCACACTTTATGATTTATCTTTGCCATTTGATGGAGTTGATGG 240  
Db 329 ACGACATTTGAAGACATCCACACATTTCTGAGTATTTCTAAGCCCTTTGTTGGATCTCTGG 388  
QY 241 AATCTCTTTTCAGCAGATGAGAAAGTTATATTAGTGGGCATAGTCTTTGGTGGTATCAAT 300  
Db 389 CGTGGCTTGTCTTAATGAAGAGTGGTCTTGTGGGCATAGCTTTGAGGAGTACAGTA 448  
QY 301 TGGGACTTGTATGGAAGATATCCACAAAGATCTATGCTGCTGTTTCTTGGCTGCTT 360  
Db 449 TAGCCCTTGCATGGACAAATCCAGAGAAATATCACTTGGAAATTTTCTTAACAGCTT 508  
QY 361 TCATGCTGATTTCTGTTACAACTCTCTCTTTGTTTGGACAGATATATGACGCGAC 420  
Db 509 TTGTTCTGTATCCCAACACAAACCATCATGCTCTTAGAAGAGTACATTTGATATACC 568  
QY 421 CAGCCGAGAAATGTTGGTACTCAGTTTTTACATATATGTTTCCCTGGAAGGCCACTGA 480  
Db 569 CATATACCGATGGATGGACACTGAGCTCTGGATAGTGGAGC-----AAAA 616  
QY 481 CATCCATGTTTTTGGCCCAAGTTCTTGGCTCACAAGCTCTACAGCTATGCTCTCCTG 540  
Db 617 CAACATTTGTTTTGGCATCAATTTCTTGTCCACTAAGTTCTATCAACTCTGCTCCACTG 676  
QY 541 AGGATCTTGCATTTAGCATCATCATTTGGTGAGACCAAGCTCTTTGTTTATGGAAGCCTAT 600  
Db 677 AGGATCTGGAATTTGTTGAAGACTTTTAAAGAAAGGGTTCACTATTGTTGTTGAAGCCTTT 736  
QY 601 CGAAGCCCAAGTATTTTCAGAGATGAACGGTTTGGATCAGTGAAGAGATTTTACATTTGCT 660  
Db 737 CTAGGAGAGAAATTTTCAAGAGAAAGATGGGTCTGTTTCCAGTGTCTATATATTT 796  
QY 661 GCACTGAGGATTAAGGCATACCAAGAAATTCAGCGATGGCAAAATGCAACATTTGGTG 720  
Db 797 CCAATGAGGACTTGGTAATTTCCAAAGAGATATCAGCAATGGATGATCCAAATGCAAGGA 856  
QY 721 TCACTGAAGCAATAGAGATTAAGGTGCTGATCAGATGCAATGCAATGCTATCGGAGCCCAAA 780  
Db 857 TTGATGTGGTGGAGAGATCAAGGGATCAGATCAGATGGTATTGCTTTAGCAAAACCCACA 916  
QY 781 AACTTTGGCCTCTCTCTTGGAAATGGCCATAA 814  
Db 917 AACTATGTTTATCTCTCTCGAGATAGCTGATAA 950

## RESULT 7

US-10-780-002-33  
; Sequence 33, Application US/10780002  
; Publication No. US20050034196A1  
; GENERAL INFORMATION:  
; APPLICANT: Klessig, Daniel F.  
; APPLICANT: Kumar, Dhirenda  
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN  
; FILE OF INVENTION: ENCODING NUCLEIC ACIDS, SAMP2, AND METHODS OF USE THEREOF  
; FILE REFERENCE: 3670-PO2652WO  
; CURRENT APPLICATION NUMBER: US/10780,002  
; PRIORITY FILING DATE: 2004-02-17  
; PRIOR APPLICATION NUMBER: PCT/US02/26312  
; PRIOR FILING DATE: 2002-08-16  
; PRIOR APPLICATION NUMBER: 60/312,863  
; PRIOR FILING DATE: 2001-08-16  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 33

; LENGTH: 870

; TYPE: DNA  
; ORGANISM: Arabidopsis Thaliana  
US-10-780-002-33

Query Match 25.4%; Score 274.6; DB 21; Length 870;  
Best Local Similarity 60.4%; Pred. No. 6.7e-67;  
Matches 489; Conservative 0; Mismatch 314; Indels 6; Gaps 2;

QY 13 AAAAGAACTAACAGGCATAAAAATTCAAATGAAGAGGAGAAAACACTTTGTTTTHAGTAC 72  
Db 1 AAAAAGTACGAAAGGAAAAATATGAGTGAAGAGAGGAAAGCAACACTTCGTGCTAGTAC 60  
QY 73 ATGTTGTCATGCCATCGAGGTTGAGTTGCTACAAGCTAAAGCCACTGCTAGAGCTGCGAG 132  
Db 61 ATGTTGCTGTCGCCGCGCATGCTGCTGCTAGAGTTAAGCTCTTCTCGAGGCTTTGG 120  
QY 133 GCCATAAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCACTGA---TTTGAGAAAAATAG 189  
Db 121 GCCATCGTGAACCGCTTTAGACCTAGCTGCTTCCGGTATACACACACAGGTCAATCA 180  
QY 190 AGAGGTTCCGACACTTTATGATTTATATCTTCCCATTTGATGGAGTTGATGGAATCTCTTT 249  
Db 181 CTGACATTTCTACATGTGAACAATATTTCTGAGCCATTTGATGAGCTAATGACTTCAATGC 240  
QY 250 CAGCAGATGAGAGGTTATATTAGTGGGCGCATAGTCTTGGTGGTATGAATTTGGGACTTG 309  
Db 241 CGAATATGAGAGGTTGACTGCTGTTGGTCATAGCTTTGGAGGTTTGAAGTTAGCCCTTAG 300  
QY 310 CTATGAAAAAGTATCCACAAAAGATCTATGCTGCTGTTTCTTGGCTGCTTTTCATGCCCTG 369  
Db 301 CCATGGAATAGTTTCCCGATAAAAATCTCTGCTCTGCTCTGCTGCTGCTGCTGCTGCTG 360  
QY 370 ATTCTGTTTCAAACTCTCTCTTTGTTTGGAAACAGTATAATGAGCGGAGCCAGCGGAGA 429  
Db 361 ACACAAAACACTCACCATCGTTCTGTCGAGGAAAGTTTGCAGAGCAGCATGACACCAAGAAG 420  
QY 430 ATTGTTTGGATCTCAGTTTTTACCATATGTTTCCCTGGAAGAGCCACTGACATCCATCT 489  
Db 421 GATGATGGGCTCTGAGCTCGAGACATATGGTT---CAGATTAATTCGGGCTTGTCTGTGT 477  
QY 490 TTTTGGGCCCAAGTTCTTTGGCTCAAAAGCTCTACAGCTATGCTCTCTGAGGATCTTTG 549  
Db 478 TCTTCAGCACCGACTTCATGAAGCACCGTCTCTACCAACTTCTCTCTGTTGAGGATCTTG 537  
QY 550 CATTAGCATCATCATTTGTTGAGACCAAGCTCTTTGTTTATGGAAGACCTATCGAAGGCCA 609  
Db 538 AGCTTGGATTTGCTTCTAAAGAGGCTAGTTCATTTGTTTAAATGAATTTATCGAAGATGG 597  
QY 610 AGTATTTTCAGAGTGAACGGTTTGGATCAGTGAAGAGAGTTTACATTTGTGCTGCTGAGG 669  
Db 598 AGAATTTTTCGAGAAAGGTTATGATCTGTTCTCGAGCTTACATTTGTGTCGAAGAGG 657  
QY 670 ATAAAGGCATACCAAGAAGATTCAGCGATGGCAAAATTCACAACTTTGTGTCTACTGAAG 729  
Db 658 ACAACATTTCTCGAAGACCATCAACGATGGATGATCCATTAATTCGGCGCATTTAG 717  
QY 730 CAATAGAGATTAAGGTTGCTGATCATGGAATGCTATGCGAGCCCCCAAAACTTTGGG 789  
Db 718 TGATTGAGATGGAAGAGACTGATCATATGCCAATGTTTGTGCAAACTCAACTACTAAGTG 777  
QY 790 CCTCTCTCTTGGAAATTTGCCCATTAATATAC 818  
Db 778 ACCATCTATTGGCAATCGCTGACAAATTC 806

## RESULT 8

US-10-780-002-34  
; Sequence 34, Application US/10780002  
; Publication No. US20050034196A1  
; GENERAL INFORMATION:  
; APPLICANT: Klessig, Daniel F.  
; APPLICANT: Kumar, Dhirenda

```
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN
; FILE REFERENCE: 3670-PO2652WO
; CURRENT APPLICATION NUMBER: US/10/780,002
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: PCT/US02/26312
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/312,863
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 837
; TYPE: DNA
; ORGANISM: Arabidopsis Thaliana
US-10-780-002-34

Query Match      25.2%; Score 271.6; DB 21; Length 837;
Best Local Similarity 59.9%; Pred. No. 4.6e-66;
Matches 492; Conservative 0; Mismatches 324; Indels 6; Gaps 2;

QY      6 GCGAAGAAAGAACTCAACAGGCATAAATTCAAATGAAGGAAGGAAAAACACTTGT 65
DB      19 GAGACAAAGTTAAAGTAGAAGAAATGAGTGAGGAGGAGGAGGACACACGTCGT 78

QY      66 TTAGTACATGTCATGCGCATCGAGGTTGGAGTTGGTACAAGCTAAAGCCACTGCTAGAA 125
DB      79 CTAGTACATGTTGCTTGCCATGGCGCTGGTCTGTGTAACAAGTTAAGCCGCGAGTCGAG 138

QY      126 GCTGCGAGGCATGAAGTTACAGCCCTTGATTTAGCAGCTTCGGCAGCTGATTTG---AGA 182
DB      139 GCTTCTGGCCACCGCGTAAACCGCGTAGATCTAGTCGCTCCGGTATAGACATGACCAAGG 198

QY      183 AAAATAGAGAGGCTTCGCACACTTTATGATTTATCTTTGCCATTCATGAGGTTGATGGAA 242
DB      199 TCATCAGACAGATATCCACATCGCGAACATYACTCAGAGCCATTGATCGAGCTAATGACC 258

QY      243 TCTCTTTCAGCAGATGAGAAGGTTATATTAGTGGGCAATGCTCTGGTGGTATGAATTTG 302
DB      259 TCACTACCAAGATGATGAGAAGGTTGTGCTTGTGGTCAATGCTTAGGAGGTTTGAGTTTA 318

QY      303 GCACTTGCTATGGAAAGTATCCACAAAGATCTATGCTGCTGTTTCTTGCTGCTGTTTC 362
DB      319 GCTATGGCCATGGATGATGTTCCGACCAAAATCTCTGTTCTGCTTGTGACTGCTATG 378

QY      363 ATGCCCTGATCTGTTCAACAATCCTCTTGTGTTTGGAAACATGATATAATGAGCGGACGCCA 422
DB      379 ATGCCAGACACCAACACACTCACCATCCTTCGTATGGATAAGCTAAGAAAAGAACTTCA 438

QY      423 GCCGAAATTTGGTGGATACTCAGTTTATCCATATGGTTCCCTGGAAGAGCCACTGACA 482
DB      439 CGAGAGGAATGGTTAGACACCGCTGTTT---ACGAGCGAAGAAACCTGATTTCTCAGCGAG 495

QY      483 TCCATGTTTTTGGCCCCAAGTTCTTGGCTCAACAGCTCTACAGCTATGCTCTCCTGAG 542
DB      496 TTTTGGATTTTGGACCAAGATTCATGCGCCAGAACTTGATCAGTTGTCTCCAGTCCAA 555

QY      543 GATCTTGCAATAGCATCATCTTGTGGACCAAGCTCTTGTGTTATGGAAGACCTATCG 602
DB      556 GATCTTGAAATGGCGAAATGTTGTGAGGGGAAACCCATTGATTAAGAAAGATATGGCA 615

QY      603 AAGGCCAAGTATTTTCAACAGTAAACGTTTGGATCAGTGAAGAGGTTTACATTTGTTGTC 662
DB      616 GAGAGAAGAGGCTTCAGTGAGGAAGGATACGGATCCGTTACACGTATATTTATTTATGTC 675

QY      663 ACTGAGGATTAAGGCATACCAAGAAATTCAGCGATGCGAAATTCAGCAACATTTGGTGTCT 722
DB      676 GGAAGAGATCTTGTGTACCCGAAGATTAACGCGATCGATGATCAGCAACTTTCGCCCA 735

QY      723 ACTGAAGCAATAGAGATTAAGAGTGTCTGATTCATGCGCAATGCTATGCGCGCCCAAAA 782
DB      736 AAGAAGTAAATGGAGATCAAGACGCGAGATCATATGCCAATGTTCTCCAAAGCTCAACA 795
```

```
QY      783 CTTTGGCGCTCTCTTGGAAATTTGCCATAAATAACAACACTGA 824
DB      796 CTATGTGCTCTTCTCTTGGAGATTGCAAAATAAATATGCTTAA 837

RESULT 9
US-09-938-842A-1551
; Sequence 1551, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krepes, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1551
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1551

Query Match      25.1%; Score 271; DB 9; Length 792;
Best Local Similarity 61.4%; Pred. No. 6.6e-66;
Matches 471; Conservative 0; Mismatches 290; Indels 6; Gaps 2;

QY      55 AACACTTGTGTTTGTAGTACATGTCATGCCATGCCATGAGGTTGGAGTTGGTACAGCTAAAGC 114
DB      23 AACACTTGTGCTGTAGTACATGTCGTGCCACGCGCATGGTGTCTGGTACAAAGGTTAAGC 82

QY      115 CACTGCTAGAAAGCTGCGAGGCCATAAGGTTACAGCCCTTGAATTTAGCAGCTTCTGGCACTG 174
DB      83 CTCTTCTCGAGCTTGTGGCCATCGTGTAAACCGCTTAGACCTAGCTGCTCTCCGTAATAG 142

QY      175 A---TTTGAGAAAAATAGAGGAGCTTCGCCACACTTTATGATATATACCTTTGGCAATGATGG 231
DB      143 ACACAACAGGTCATCACTGACATTTCTACATGTGAACAATATTTCTGAGCCATTTGATGC 202

QY      232 AGTTGATGGAATCTTTTACAGCAGATGAGAAGGTTATATATTAGTGGGGCATAGTCTTGGTG 291
DB      203 AGCTAATGACTTCATTGCGCGAATGATGAGAAGGTTGTACTCGTTGGTCATAGCTTTGGAG 262

QY      292 GTATGAAATTTGGGACTTCTGCTATGAAAAAGTATCCAAAAAGATCTATCTGCTGCTTTTCT 351
DB      263 GTTTGAGTTTAGCTTACCATGATAGTTTCCGATAAATCTCTGCTCTGCTCTGCTTTCG 322

QY      352 TGGCTGCTTTTCATGCTGATTTCTGTTCAACAATCTCTCTCTTTGTTTGGAAACAGTATAATG 411
DB      323 TGACTGCAATTCATGCCCCGACCAACAACACTCACCATCTGTCGAGGAAAGTTTGCAG 382

QY      412 AGCGAGCCCGAGCCGAGAAATTTGGTGGATACCTCAGTTTTTACCATATGTTTCCCTGAG 471
DB      383 CGAGCATGACACCCAGAGAGGATGGATGGGCTCTGAGCTCGAGACATATGTTT---CAGATA 439

QY      472 AGCCACTGACATCATGTTTGTGGCCCAAAAGTTCTTGGCTCACAAGCTCTACCAAGCTAT 531
DB      440 ATTCCGGCTTGTCTGTTCTTTCAGCACCGACTTCATGAAGCACCGTCTCTACCAACTTT 499

QY      532 GCTCTCTGAGAGATCTTCATTTAGCATCATCAITGGTGGAGCAAGCTCTTTGTTTATGG 591
DB      500 CTCCTGTGGAGATCTTGAGCTTGGATTTGCTTTTAAAGAGGCGCTAGTTCAITGTTTATTA 559

QY      592 AAGACCTATCGAAGGCCAAGTATTTTTCACAGATGAACGGTTTGGATCACTGAGGAGAGTTT 651
```



Db 560 ATGAATATCGAAGATCGAGAACTTTCTGAGAAAGGATATGATCTGTTCTCGAGCTT 619  
Qy 652 ACATTGTGTGCACTGAGGTAAGGCATACCGAGAAATTCAGCGATGCGAAATTGACA 711  
Db 620 ACATTGTGTGCAAGAGGACAACTTATCTCGAAGACCATCAACGATCGATGCCATA 679  
Qy 712 ACATTGGTGTCACTGAAGCAATAGAGATTAAAGGTCTGATCACATGCGCAATGCTATGG 771  
Db 680 ATTATCCGGGATTTAGTGTAGATGGAAGAGACTGATCATATGCCAAATGTTTGCA 739  
Qy 772 AGCCCCAAAACCTTTGGCGCTCTCTCTGGAATTCGCCATTAATAC 818  
Db 740 AACCTCAACTACTAAGTGACCATCTATTGCAATCTGACCAATTC 786

## RESULT 10

US-09-938-842A-1551  
; Sequence 1551, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRIPT300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1551  
; LENGTH: 792  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1551

Query Match 25.1%; Score 271; DB 11; Length 792;  
Best Local Similarity 61.4%; Pred. No. 6.6e-66;  
Matches 471; Conservative 0; Mismatches 290; Indels 6; Gaps 2;

Qy 55 AACACTTTGTTTGTAGTACATGGTGCATGCCATGGAGTTGGAGTTGGTACAAAGCTAAAGC 114  
Db 23 AACACTTTGTTGCTAGTACATGGTGCATGGTGCACGGCGCATGGTGTGTGTTAAAGCTAAAGC 82  
Qy 115 CACTGCTAGAGCTGCAGGCCATAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCACTG 174  
Db 83 CTCCTTCGAGGCTTTGGGCCCATCGTGTAAACGCCCTTAGACCTAGCTGCTCCGGTATAG 142  
Qy 175 A---TTTGAGAAAATAGAGGAGCTTCGCACACTTTATGATATATCTTTGCCATTGATGG 231  
Db 143 ACACACACGAGTCAATCACTGACATTTCTACATGTGCAACAATATCTGAGGCCATTGATGC 202  
Qy 232 AGTTGATGAAATCTCTTTTCAGCAGATGAGAAGTTATATTAGTGGGGCATAGTCTTGGTG 291  
Db 203 AGCTAATGACTTCATTGCGCAATGATGAGAAGTTGTACTCGTTGGTCATAGCTTTGGAG 262  
Qy 292 GTATGAAATTTGGACTTGTCTATGGAAGATATCCAAAGATCTATGCTGCTGTTTCT 351  
Db 263 GTTTGAGTTTAGCTTTAGCCATGGATTAAGTTTCCCGATAAAATCTGTCTCTGCTTTCG 322  
Qy 352 TGGCTCTTTTCATGCTGATTTCTGTTCAACAATCTCTCTTTGTTTGGAAACAGTATAATG 411  
Db 323 TGACTGCATTTATGCCCCGACACCAACACTCACCATCTGTTCTGCGAGGAAAATTTGCA 382  
Qy 412 AGCGGACGCCAGCCGAGAAATTTGGTTGGATACTCAGTTTTTACCATATGTTCCCTCGAAG 471

Db 383 GCAGCATGACACACAGAGGATGATGGGCTCTGAGCTCGAGACATATGGTT---CAGATA 439  
Qy 472 AGCCACTGACATCCATGTTTTTTGGCCCAAAAGTTCTTGGCTCACAAGCTCTACCAAGCTAT 531  
Db 440 ATTCCGGCTTGTCTGTGTTCTTCAGACCGACTTCATGAGCCCGCTCTCTACCAACTTT 499  
Qy 532 GCTCTCTCGAGGATCTTGCATTTAGCATCATCTTGGTGAGACCAAGCTCTTTGTTTATGG 591  
Db 500 CTCTCTGTGGAGGATCTTGAGCTTGGATTGCTTCTAAAGAGGCTAGTTTCATTGTTTATTA 559  
Qy 592 AAGACTATCGAAGCCCAAGTATTTTCAGATGAACGGTTTGGATCAGTGAAGAGATTT 651  
Db 560 ATGAATTTATCGAAGATGGAGAACTTTTCTGAGAAAGGGTATGGATCTGTTCCTCGAGCTT 619  
Qy 652 ACATTGTGTGCACTGAGGATAAGGCATACAGAGAAGAAATTCAGCGATGGCAAAATTGACA 711  
Db 620 ACATTGTGTGCAAGAGGACAACTTATCTCGAAGACCATCAACGATGGATGCCATA 679  
Qy 712 ACATTGGTGTCACTGAAGCAATAGAGATTAAAGGTGCTGATCACATGGCAATGCTATGG 771  
Db 680 ATTATCCGGCGAAATTTAGTGTAGATGGAAGAGACTGATCATATGCCAAATGTTTGCA 739  
Qy 772 AGCCCCAAAACCTTTGGCGCTCTCTCTTGGAAAATTTGCCATTAATAC 818  
Db 740 AACCTCAACTACTAAGTGACCATCTATTGGAATTCGCTGACCAATTC 786

## RESULT 11

US-09-938-842A-217  
; Sequence 217, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRIPT300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 217  
; LENGTH: 792  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-217

Query Match 24.9%; Score 269; DB 9; Length 792;  
Best Local Similarity 61.1%; Pred. No. 2.4e-65;  
Matches 472; Conservative 0; Mismatches 295; Indels 6; Gaps 2;

Qy 55 AACACTTTGTTTGTAGTACATGGTGCATGCCATGGAGGTTGGAGTTGGTACAAAGCTAAAGC 114  
Db 23 AACAGTCGTCTAGTACATGGTGTCTTGCCATGGCGCCTGGTGTGTGTTACAAAGGTTAAGC 82  
Qy 115 CACTGCTAGAGCTGCAGGCCATAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCACTG 174  
Db 83 CGCAGCTCGAGGCTTCTGGCCACCGCGTAACCGCGTAGATCTAGCTGCTCCGGTATAG 142  
Qy 175 ATTTG---AGAAAATAGAGGAGCTTCGCACACTTTTATGATTTATCTTTGCCATTGATGG 231  
Db 143 ACATGACAGGTCATATCAGATATATCCACATGCGAACAATACTCAGAGCCATTGATGC 202  
Qy 232 AGTTGATGGAATCTCTTTTCAGCAGATGAGAAGGTTATATTAGTGGGGCATAGTCTTGGTG 291  
Db 203 AGCTAATGACCTCACTACCAAGATGATGAGAAGGTTGTGCTTGTGTTGTTGTTAGGAG 262

```
QY 292 GTATGAATTTGGGACTTGTCTATGGAAGATATCCAAAGATCTATGCTGCTGTTTCT 351
Db 263 GTTTCAGTTTATGCTATGGCCATGATGTTTCCGACCAAAATCTCTGTTCTGTCTTTG 322
QY 352 TGGCTGCTTTATGCTGATTTCTGTTTCAAACTCTCTCTTGTGTTGGAACAGTAAATG 411
Db 323 TGACTGCTATGATGCGAGACACAAACACTCAACCTTCTGTATGGGATAAGCTAAAGAA 382
QY 412 AGCGGACGCGCAGCGAGAAATTTGGTGGATPACTCAGTTTTTACCATATATGTTCCCTGGAAG 471
Db 383 AAGAACTTTCAGAGGGAATGTTTAGACACCGTGT---ACGAGCGAGAAACCTGATT 439
QY 472 AGCCACTGACATCCATGTTTTTTTGGCCCAAAAGTTCTTGCTGCTCACAAGCTCTACCAAGCTAT 531
Db 440 TTCTAGGAGGTTTTGGATTTTGGACCAAGAAATTCATGGCCAAAGAACTTGTATCAGTTGT 499
QY 532 GCTCTCCTGAGGATCTTGATTTAGCATCATCTGTTGAGACCAAGCTCTTTGTTTATGG 591
Db 500 CTCCAGTCCAAATCTTGAATTTGGCGAAATGTTTGGTGAAGGCAACCCATTTGATTAAGA 559
QY 592 AAGACCTATCGAAGGCCAAGTATTTTCCAGATGAACGGTTTGGATCAGTGAAGAGAGTTT 651
Db 560 AAGATATGCGAGAGAGAAAGCTTCAGTGAGAGGATACGGATCCGTTACACGTATAT 619
QY 652 ACATTGTGTGCTGAGGATTAAGGCAATACCAAGAAATTCAGCGATGGCAAAATTGACA 711
Db 620 TTATTGTATGCGAAAGGATCTTGTGTCAACCGAAGATTTACCAGCGATCGATGATCAGCA 679
QY 712 ACATTGTGTCTGAGCAATAGAGATTAAGAGTTCATGATCAGATGCAATGCTATGCG 771
Db 680 ACTTTCCCAAGAAAGTAAATGAGATCAAGAGCGAGATCATATGCCAATGTTCTCCA 739
QY 772 AGCCCCAAAAATTTTGGCGCTCTCTCTTGGAAATTTGCCATATAAATACAACTCA 824
Db 740 AGCCTCAACAATATGTGCTCTCTCTTGGAGATTGCAATTAATATGCTTAA 792
```

## RESULT 12

```
US-09-938-842A-217
; Sequence 217, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 217
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-217
```

```
Query Match 24.9%; Score 269; DB 11; Length 792;
Best Local Similarity 61.1%; Pred. No. 2.4e-65;
Matches 472; Conservative 0; Mismatches 295; Indels 6; Gaps 2;
QY 55 AACACTTGTGTTTAGTACATGGTGCATGCCATGGAGTTGGTACAGCTAAAGC 114
Db 23 AACACGTCGTTTGTAGTACATGGTGTGCTTGCATGGCGCTGTGTTGATCAAGGTTAAGC 82
```

```
QY 115 CACTGCTAGAGCTGCGAGGCCATAAGGTTACAGCCCTTGATTTTAGAGCTCTTGGCACTG 174
Db 83 CGCAGCTGAGGCTTCTGCGCCACCGCGTAACCGCGTAGATCTAGCTGCTCCGTTAGTAG 142
QY 175 ATTTG---AGAAAAATAGAGAGCTTGGCACACATTTTATGATTAATCTTTGGCATTTGATGG 231
Db 143 ACATGACCAAGTCAATCACAGATATATCCACATCGGAACAATACTACAGAGCAATTGATGC 202
QY 232 AGTTGATGAATCTCTTTTACGACATGAGAAAGTTATATAGTGGGGCATAGTCTTGGTG 291
Db 203 AGCTAAATGACCTCACTACCAAGATGATGAGAAAGTTGTCTTGTGGTTCATAGCTTAGGAG 262
QY 292 GTATGAATTTGGGACTTGTCTATGGAAGATATCCAAAAAGATCTATGCTCCTGTTTCT 351
Db 263 GTTTGAGTTTATGCTATGGCCATGATATGTTTCCGACCAAAATCTCTGTTTCTGTCTTTG 322
QY 352 TGGCTGCTTTTCAATGCTGATTTCTGTTTCAAACTCCTCTCTTTTGGTGGAAACAGTATAATG 411
Db 323 TGACTGCTATGATGCCAGACACCAAACTCACTCAACATCCTTCGTATGGGATAAGCTAAAGAA 382
QY 412 AGCGGACGCCAGCGAGAAATTTGGTGGATATCTCAGTTTATACCATATATGGTTCCCTGAA 471
Db 383 AAGAACTTTCAGGAGAGGAATGTTTAGACACCGTGT---ACGAGCGAGAAACCTGATT 439
QY 472 AGCCACTGACATCCATGTTTTCGCCCAAAAGTTTCTTGGCTCACAAGCTCTACCAAGCTAT 531
Db 440 TTCTTAGCGAGTTTGGGATTTTGGACCAAGAAATTCATGGCCAAAGAACTTGTATCAGTTGT 499
QY 532 GCTCTCCTGAGGATCTTTCATTTAGCATCATCAATTTGGTGAGACCAAGCTCTTTGTTTATGG 591
Db 500 CTCCAGTCCAAAGATCTTGAATTTGGGAAATGTTTGGTGAAGGCAACCCATTTGATTAAGA 559
QY 592 AAGACTATCGAAGGCCAAGTATTTTCAAGATGAACGGTTTGGATCAGTGAAGAGAGTTT 651
Db 560 AAGATATGCGAGAGAGAAAGCTTCAGTGAGGAGGATACCGATCCGTTACAGTATAT 619
QY 652 ACATTGTGTGCACTGAGGATAAAGCATACCAAGAAATTCAGCGATGGCAAAATTGACA 711
Db 620 TTATTGTATGCGAAAGGATCTTGTGTACCCGGAAGATTACAGCGATCGATGATCAGCA 679
QY 712 ACATTGTGTCTCACTGAAGCAATAGAGATTAAGGTTGCTGATCACAATGGCAATGCTATCGG 771
Db 680 ACTTTCCCAAGAAAGTAAATGAGATCAAGAGCGAGATCATATGCCAATGTTCTCCA 739
QY 772 AGCCCCAAAAATTTTGGCGCTCTCTCTTGGAAATTTGCCATATAAATACAACTCA 824
Db 740 AGCCTCAACAATATGTGCTCTCTCTTGGAGATTGCAAAATAAATATGCTTAA 792
```

## RESULT 13

```
US-10-425-114-12800
; Sequence 12800, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kowalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 12800
; LENGTH: 1035
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 701212327_FLI
US-10-425-114-12800
```



RESULT 15  
US-10-780-002-37  
; Sequence 37, Application US/10780002  
; Publication No. US20050034196A1  
; GENERAL INFORMATION:  
; APPLICANT: Klessig, Daniel F.  
; APPLICANT: Kumar, Dhirenda  
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN  
; TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, SAMP2, AND METHODS OF USE THEREOF  
; FILE REFERENCE: 3670-P02652WO  
; CURRENT APPLICATION NUMBER: US/10/780,002  
; CURRENT FILING DATE: 2004-02-17  
; PRIOR APPLICATION NUMBER: PCT/US02/26312  
; PRIOR FILING DATE: 2002-08-16  
; PRIOR APPLICATION NUMBER: 60/312,863  
; PRIOR FILING DATE: 2001-08-16  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 37  
; TYPE: DNA  
; ORGANISM: Arabidopsis Thaliana  
US-10-780-002-37

Query Match 23.4%; Score 252.6; DB 21; Length 819;  
Best Local Similarity 58.2%; Pred. No. 1.1e-60;  
Matches 463; Conservative 0; Mismatches 329; Indels 3; Gaps 1;

|    |     |                |              |            |        |            |           |               |     |
|----|-----|----------------|--------------|------------|--------|------------|-----------|---------------|-----|
| QY | 26  | RAGGCATAAAATTC | AAATCAAGGAGG | AGGAAACAC  | TTTGT  | TTTAGTACAT | GGTG      | GCATGCCA      | 85  |
| DB | 12  | AAGAAGGCAAT    | TGGGAATAGAA  | CCAGAACG   | GT     | TTGTCTCAT  | CCAT      | GGGGTGTGCCA   | 71  |
| QY | 86  | TGGAGGTTG      | CGAGTTGGTCA  | AGCTAAAGCC | CACTG  | CTAGAA     | GCTGCAGG  | CCCATAAAGTTAC | 145 |
| DB | 72  | CGGGCATG       | GCACCTTGGG   | CAAGGTG    | AAACG  | CAGCTGG    | AGGTTG    | CAGGTCAC      | 131 |
| QY | 146 | AGCCCTTG       | ATTTAGCAG    | CTTCTGG    | CAC    | TGATTTG    | TGA       | AAAAATAGAGG   | 205 |
| DB | 132 | GCAGTGG        | ATCTTGTG     | CATCAG     | GTAT   | AAACAT     | ACG       | CAAAAGT       | 191 |
| QY | 206 | TTATGAT        | TATAC        | TTTGG      | CAATTC | GAGTTG     | ATG       | GAATCTCTT     | 262 |
| DB | 192 | GAA            | CGAATTA      | CTGCA      | AAACCA | TGCTTG     | AGTTT     | CTGAGCTCG     | 251 |
| QY | 263 | GGTTAT         | TATTTAGTGGG  | CATAGT     | CTTTGG | TGGTAT     | GAAATTTGG | GACTTGGCTATG  | 322 |
| DB | 252 | GGTGAT         | TGTTGTTG     | CTCAT      | TAGCAT | GGAGG      | AAATAT    | CCGCTGC       | 311 |
| QY | 323 | TCACAAA        | AGATCTAT     | GTGCTG     | TTTCTT | GGCTG      | CTTTC     | ATGCTGTT      | 382 |
| DB | 312 | CGCTTG         | TAAAGATTG    | CCGCTAT    | TGCTCT | TTTGA      | CAGCTTT   | CATGCC        | 371 |
| QY | 383 | CTCCTC         | CTTTGTTT     | GGACAG     | TATAT  | AGCGG      | AGCCG     | CAGAGATTTG    | 442 |
| DB | 372 | ACCTG          | CTTAATGTT    | TACGAA     | AGCTG  | CTCAG      | AAGCAT    | TTCCAA        | 431 |
| QY | 443 | TCAGT          | TTTTTAC      | CAATAT     | TGTTCC | CTGAG      | AGCCACT   | GCATCCAT      | 502 |
| DB | 432 | CACG           | TGTGTGA      | CTAC       | CGGGA  | ACCTG      | ATTTT     | CTCTAC        | 491 |
| QY | 503 | GTTCTT         | GGCTCA       | CAAGCT     | CTAC   | CAGCTAT    | GCTCT     | CTCCTG        | 562 |
| DB | 492 | GTTTAT         | GGCCAA       | AAAAAT     | GTAT   | CAAA       | AACTCT    | CCAGTT        | 551 |
| QY | 563 | ATTG           | TGAGCA       | CAAGCT     | CTTTG  | TTATG      | GAGAC     | CTTATCG       | 622 |
| DB | 552 | ATTAG          | TGAGG        | AAAAAC     | CCCGT  | TAGTT      | ACAA      | CAATCT        | 611 |
| QY | 623 | TGA            | CCGTTT       | TGGAT      | CAGTGA | AGAGT      | TTTAC     | ATTG          | 682 |
| DB | 612 | GGA            | AGGGT        | ACGGAT     | CCGTT  | AC         | GTTAT     | TATAT         | 671 |

Search completed: September 12, 2005, 12:48:34  
Job time : 2393.95 secs

|    |     |        |        |       |      |      |     |    |      |     |      |     |    |    |    |     |    |     |    |     |     |
|----|-----|--------|--------|-------|------|------|-----|----|------|-----|------|-----|----|----|----|-----|----|-----|----|-----|-----|
| QY | 683 | AGAGAA | TTCCAG | CATGG | CAAA | TTG  | CA  | CA | TTGG | TGT | CT   | CA  | CT | GA | GC | AA  | TA | GAG | AT | TAA | 742 |
| DB | 672 | CGA    | AGAT   | TACC  | AGCG | TTGG | ATG | AT | AG   | CA  | CTTT | CC  | AC | CA | AA | AG  | AT | AT  | GG | AT  | 731 |
| QY | 743 | AGGTG  | CTG    | AT    | CA   | CA   | TG  | CA | AT   | G   | CT   | TAT | G  | C  | G  | A   | C  | C   | C  | C   | 802 |
| DB | 732 | ATG    | TG     | CA    | AT   | CA   | TAT | G  | CA   | AT  | G    | TT  | CT | CA | A  | G   | C  | C   | T  | CT  | 791 |
| QY | 803 | AA     | TTG    | CC    | CA   | TA   | AA  | TA | A    | 817 |      |     |    |    |    |     |    |     |    |     |     |
| DB | 792 | G      | ATT    | G     | CA   | AA   | T   | A  | A    | T   | A    | A   | T  | A  | A  | 806 |    |     |    |     |     |

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 01:44:06 ; Search time 173.605 Seconds  
(without alignments)  
8595.849 Million cell updates/sec

Title: US-10-780-002-36

Perfect score: 912

Sequence: 1 atcagctattctcaacaa.....ttttccgggcaactttcatc 912

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/1/ina/5A COMB.seq:\*
- 2: /cgn2\_6/prodata/1/ina/5B COMB.seq:\*
- 3: /cgn2\_6/prodata/1/ina/6A COMB.seq:\*
- 4: /cgn2\_6/prodata/1/ina/6B COMB.seq:\*
- 5: /cgn2\_6/prodata/1/ina/PTUS COMB.seq:\*
- 6: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description          |
|------------|-------|-------------|--------|----|----------------------|
| 1          | 221.6 | 24.3        | 1078   | 3  | US-08-981-256A-4     |
| 2          | 43    | 4.7         | 7218   | 1  | US-08-232-463-14     |
| 3          | 38.6  | 4.2         | 1227   | 4  | US-09-252-991A-13741 |
| 4          | 38.6  | 4.2         | 1473   | 4  | US-09-252-991A-13404 |
| 5          | 36.2  | 4.0         | 640681 | 4  | US-09-790-988-1      |
| 6          | 35.4  | 3.9         | 459    | 3  | US-09-427-700-8      |
| 7          | 35.4  | 3.9         | 572    | 1  | US-07-989-363-1      |
| 8          | 35.4  | 3.9         | 572    | 1  | US-08-264-526-1      |
| 9          | 35.4  | 3.9         | 1442   | 1  | US-08-152-483B-8     |
| 10         | 35.4  | 3.9         | 4009   | 1  | US-08-152-483B-6     |
| 11         | 35.4  | 3.9         | 4294   | 1  | US-08-152-483B-2     |
| 12         | 35.4  | 3.9         | 42094  | 4  | US-09-949-016-16445  |
| 13         | 34.8  | 3.8         | 832    | 4  | US-09-621-976-2813   |
| 14         | 34.8  | 3.8         | 175265 | 4  | US-09-949-016-16089  |
| 15         | 34.4  | 3.8         | 601    | 4  | US-09-949-016-57364  |
| 16         | 34.4  | 3.8         | 49378  | 4  | US-09-949-016-13408  |
| 17         | 34.4  | 3.8         | 317366 | 4  | US-09-949-016-16001  |
| 18         | 34    | 3.7         | 312474 | 4  | US-09-949-016-17434  |
| 19         | 33.8  | 3.7         | 154023 | 4  | US-09-949-016-17057  |
| 20         | 33.6  | 3.7         | 1141   | 4  | US-09-806-708B-22    |
| 21         | 33.6  | 3.7         | 1750   | 3  | US-08-949-770-1      |
| 22         | 33.6  | 3.7         | 26763  | 4  | US-09-949-016-11984  |
| 23         | 33.6  | 3.7         | 26764  | 4  | US-09-949-016-15566  |
| 24         | 33.4  | 3.7         | 1122   | 4  | US-09-252-991A-11137 |
| 25         | 33.4  | 3.7         | 1260   | 4  | US-09-252-991A-10925 |
| 26         | 33.2  | 3.6         | 601    | 4  | US-09-949-016-46899  |
| 27         | 33.2  | 3.6         | 747    | 4  | US-09-252-991A-3590  |

|      |      |     |        |   |                     |                   |
|------|------|-----|--------|---|---------------------|-------------------|
| c 28 | 33.2 | 3.6 | 1191   | 4 | US-09-252-991A-3692 | Sequence 3692, Ap |
| c 29 | 33.2 | 3.6 | 1845   | 4 | US-09-252-991A-3639 | Sequence 3639, Ap |
| c 30 | 33.2 | 3.6 | 246444 | 4 | US-09-949-016-13113 | Sequence 13113, A |
| c 31 | 33   | 3.6 | 2525   | 3 | US-09-342-648-1     | Sequence 1, Appli |
| c 32 | 33   | 3.6 | 26354  | 4 | US-09-949-016-12746 | Sequence 12746, A |
| c 33 | 33   | 3.6 | 26355  | 4 | US-09-949-016-12995 | Sequence 12995, A |
| c 34 | 33   | 3.6 | 140925 | 4 | US-09-949-016-11777 | Sequence 11777, A |
| c 35 | 33   | 3.6 | 140982 | 4 | US-09-949-016-16295 | Sequence 16295, A |
| c 36 | 32.8 | 3.6 | 601    | 4 | US-09-949-016-79644 | Sequence 79644, A |
| c 37 | 32.8 | 3.6 | 807    | 4 | US-09-270-767-13804 | Sequence 13804, A |
| c 38 | 32.8 | 3.6 | 1500   | 1 | US-08-476-008-6     | Sequence 6, Appli |
| c 39 | 32.8 | 3.6 | 1500   | 1 | US-08-306-063-6     | Sequence 6, Appli |
| c 40 | 32.8 | 3.6 | 1500   | 1 | US-08-833-485-6     | Sequence 6, Appli |
| c 41 | 32.8 | 3.6 | 1500   | 3 | US-09-137-440-6     | Sequence 6, Appli |
| c 42 | 32.8 | 3.6 | 1500   | 5 | PCT-US91-06148A-6   | Sequence 6, Appli |
| c 43 | 32.8 | 3.6 | 1673   | 1 | US-08-476-008-4     | Sequence 4, Appli |
| c 44 | 32.8 | 3.6 | 1673   | 1 | US-08-306-063-4     | Sequence 4, Appli |
| c 45 | 32.8 | 3.6 | 1673   | 1 | US-08-833-485-4     | Sequence 4, Appli |

ALIGNMENTS

RESULT 1

US-08-981-256A-4  
; Sequence 4, Application US/08981256A  
; Patent No. 6046042  
; GENERAL INFORMATION:  
; APPLICANT: Meinhard HASSLACHER et al.  
; TITLE OF INVENTION: (S)-HYDROXY-NITRILE-LYASE FROM HEVEA BRASILIENSIS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.  
; STREET: 2033 K Street, N.W., Suite 800  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/981.256A  
; FILING DATE: December 22, 1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew Jacob  
; REGISTRATION NUMBER: 25,154  
; REFERENCE/DOCKET NUMBER: 1553-021112  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 721-8200  
; TELEFAX: (202) 721-8250  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1078 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-981-256A-4

|                       |        |   |          |            |        |        |    |      |    |
|-----------------------|--------|---|----------|------------|--------|--------|----|------|----|
| Query Match           | 24.3%; | Score   | 221.6;   | DB 3;      | Length | 1078;  |    |      |    |
| Best Local Similarity | 55.3%; | Pred. NO.   | 6.2e-64; |            |        |        |    |      |    |
| Matches               | 431;   | Conservative  | 0;       | Mismatches | 349;   | Indels | 0; | Gaps | 0; |
| QY                    | 76     | TCACCTTCGTGTAGTTCACACGCTTATCATGGAGCCTGGATCTCGGTCAAGCTCAAGCC | 135      |            |        |        |    |      |    |
| Db                    | 55     | TCATTTGTTCTTATTTCATATATGCGCGGTGCTGATTTTGGCACAGCTCAAAACC     | 114      |            |        |        |    |      |    |

```
QY 136 CCTCTCTGAAATCAGCGCGGCCACCGCGTTACTCTCTCGAATCTCGCGCGCTCGGAGTCGA 195
Db 115 CCTCTCTGAGGCACTTGGCGCAAGGTTACTCACTGGACCTTGCGCAAGCGCGTTGA 174
QY 196 CCACGACCAATCCAGCGCGTTGAAACCGTCGACCAATACTCCAAACCGTTGATCGAATC 255
Db 175 CCCAAGGCANAATTGAGGAGATTGGCTCATTTGATGAGTATCTGAAACCTTGTTCAGCTT 234
QY 256 CCTCAAAATCTTCCAGAGAAACGAGAGATTAATCTGGTTGGATTTCAGCTTCGGAGGAT 315
Db 235 CTTGGAGGCACTCCCTCCAGGGGAAAGGTGATCTGGTTGGCGAGAGCTGTGGAGACT 294
QY 316 CAACATCGCTCTCGCGCGCAATATTTCCCGCGAAGATAAAGGTTCTTGTGTTCTCTCAA 375
Db 295 CAATATAGCAATTCGTCTGTATAAATACTGTGAAAAGATTGCAGCTGTCTTTTCCACAA 354
QY 376 CGCCTTCTGCGGACACCAACCGTCCTCTCAGCTTCGGAAGATATATCGAGAT 435
Db 355 TTCAGTATTGGCAGACACCGACACTGCCCATCTTACGTCGGGATAAGCTCATGGAGGT 414
QY 436 GCGTCGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAACAAAGAAATGGGACGATGAG 495
Db 415 GTTCCCGACTGGAAGACACACAGTATTTACGTACACTAAAGATGGCAAGGAGATAAC 474
QY 496 TTTATTGAAGATGGGACCAAAATTCATGAAGCAGCTCTTTACCAAAATTTGTCCTATGA 555
Db 475 TGGATTGAACTGGGCTTCACGCTTCTGAGGGAAAATTTATATACCTTTTGGCGTCTGA 534
QY 556 GGAATTACGAGTCGCAAAATCTGCATAGCAAGGTCATTTTTCAGAGGATCTATC 615
Db 535 GGAATATGAATCTGGCAAGATGTTGCAAGGAAGGATCATTTTCAAAAATATTTTACG 594
QY 616 AAGAAAGAAAATTTAGCGGAAAGGATATGTTTCGGTGCACACAGTTTACGTAATGAG 675
Db 595 TAAGCGACCATCTTCACTAAGGAGTTACGATCGATTAGAAAATTTATGTGTGAC 654
QY 676 TAGTGAAGCAAAAGCCATCCCTGCGATTTCATTCGTTGGATGATGATAATTTCAACGT 735
Db 655 CGACCAAGACGAAATATTTTACCTGAAATTTCAACTCTGGCAATAGAAAACATATAAAC 714
QY 736 CTCGAAGCTACGAGATCGATGGCGGAGATCACATGTTGATGCTCCAAACCCCAAA 795
Db 715 AGACAAGGTTTAAAGTTCGAAGGTGGAGATCATAAATTCGACCTTACAAAGACTAAGA 774
QY 796 ACTCTTTGACTCTCTCTCTGCTATTGCCCACCGATTATATGTAATAATCTTAAGTCCGTT 855
Db 775 GATCGCTGAATCTCCAGAGGTGGCTGATACCTATAATTGACTCTTTTGAGGCTTTT 834
```

## RESULT 2

```
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 INMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
; US-08-232-463-14
Query Match 4.7%; Score 43; DB 1; Length 7218;
Best Local Similarity 3.5%; Pred. No. 0.0065;
Matches 10; Conservative 165; Mismatches 110; Indels 0; Gaps 0;
QY 418 GGACAAGTATATGAGATGCTCGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAC 477
Db 1334 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1275
QY 478 AAGAAATGGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGCACGCTCTTTA 537
Db 1274 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1215
QY 538 CCAAAATTTGCCATAGAGGATTACGAGCTGGCAAAATTTGTCATAGGCAAGGTCATT 597
Db 1214 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1155
QY 598 TTTCCAGAGATCTATCAAGAAAGAAAGTTTAGCGGAGGAGATATGTTTCGGTGA 657
Db 1154 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1095
QY 658 ACAGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTGCGA 702
Db 1094 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1050
```

## RESULT 3

```
US-09-252-991A-13741/c
; Sequence 13741, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13741
; LENGTH: 1227
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-13741
```

```
Query Match      4.2%; Score 38.6; DB 4; Length 1227;
Best Local Similarity 54.6%; Pred. No. 0.056;
Matches 77; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 208 CCAGGCGGTTGAACCGTCGACGAATACTCCAACCGTTGATCGAAACCCCTCAATCTCT 267
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 442 CCAGCGGTGGCCGCGATCGACCACTACCGGACCTGTTGCTCGACACCCCTCAAGCGGC 383
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 268 TCCAGAGAAGAGAGGTAAATTCGTGTGATTGAGTTTCGGAGGATCAACATCGCTCT 327
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 382 CAGTGACGTGATCGCGAGCGGTGTGATTGACCGCGGACGTTTCAACAGCGCCTA 323
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 328 CGCGCGCGACATATTTCCGCG 348
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 322 CTTCGAACACGCGCTTCTCGC 302
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
US-09-252-991A-13404
; Sequence 13404, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13404
; LENGTH: 1473
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13404

Query Match      4.2%; Score 38.6; DB 4; Length 1473;
Best Local Similarity 54.8%; Pred. No. 0.064;
Matches 77; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 208 CCAGGCGGTTGAACCGTCGACGAATACTCCAACCGTTGATCGAAACCCCTCAATCTCT 267
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 621 CCAGCGGTGGCCGCGATCGACCACTACCGGACCTGTTGCTCGACACCCCTCAAGCGGC 680
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 268 TCAGAGAACGAAGAGGTAAATTCGTGTGATTGAGTTTCGGAGGATCAACATCGCTCT 327
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 681 CAGTGACGTGGAATCGCGAGCGTGTGATTGACCGCGGACGTTTCAACAGCGCCTA 740
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 328 CGCGCGCGACATATTTCCGCG 348
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 741 CTTCGAACACGCGCTTCTCGC 761
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEKI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match      4.0%; Score 36.2; DB 4; Length 640681;
Best Local Similarity 55.0%; Pred. No. 32;
Matches 71; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 540 AAAATTGTCCTATAGAGGATTACGAGCTGGCAAAAATGTTGCATAGGCAAGGTCATTTT 599
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 59751 AAAATTAAATTAATAATAGAGAAAGAACTGGCTTAACTAATTGAACAAGTAAAAAGATATTAA 59692
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 600 TCACAGAGGATCTATCAAAAGAAAGAAAAGTTTTAGCGAGGAAGATATGTTCCGTCGAAC 659
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 59691 TAAAGATGTCCTATTGGAGAGCAAGCTAGACGCGGAAAAAGAAATGTCGAACG 59632
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 660 GAGTTTACG 668
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 59631 TAAATTACG 59623
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
US-09-427-700-8/c
; Sequence 8, Application US/09427700
; Patent No. 6372489
; GENERAL INFORMATION:
; APPLICANT: Anticancer
; APPLICANT: Zhao, Ming
; TITLE OF INVENTION: METHOD AND MODEL FOR HAIR PIGMENTATION
; FILE REFERENCE: 31276-20017.00
; CURRENT APPLICATION NUMBER: US/09/427,700
; CURRENT FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: US 60/105,725
; PRIOR FILING DATE: 1998-10-27
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 459
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (16)...(450)
; OTHER INFORMATION: ORF 438 gene
US-09-427-700-8

Query Match      3.9%; Score 35.4; DB 3; Length 459;
Best Local Similarity 51.6%; Pred. No. 0.34;
Matches 81; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 74 CATCACTCGTGTAGTTTCACACCGCTTATCATATGAGCGCTGATCTGGTACAGCTCAAG 133
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 375 CACCGGCTCGTAGTGGCTGACGACGCTGATCCACGAGCGCTGCGGCTTCCGATCACATG 316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 134 CCCCTCCTTGAATCAGCGCGCCACCGCTTACTGCTGCAACTCGCCGCTCCGGGATC 193
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 315 CAGTTGAGCGCGCTCCACGACACGCGGTAAACCGCGCGGTGATGCGCCCGCGTGACC 256
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 194 GACCCACGACCAATCAGGCGGTTGAAACCGTCGACG 230
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 255 GTCACCGCGTGACGCGCGCTGTGTGGCCCCCGCG 219
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
US-07-989-363-1/c
; Sequence 1, Application US/07989363
; Patent No. 5385841
; GENERAL INFORMATION:
; APPLICANT: Ortega, Jose Daza
; APPLICANT: Gill, Jose Antonio
; APPLICANT: Garcia, Tomas Vigal
```







```
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/152,483B
/ FILING DATE: No. 5529909ember 12, 1993
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 7/857,602
/ FILING DATE: March 30, 1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 923,692
/ FILING DATE: July 31, 1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 600,244
/ FILING DATE: October 22, 1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 641,617
/ FILING DATE: January 16, 1991
/ NAME: Halluin, Albert P.
/ REGISTRATION NUMBER: 25,227
/ REFERENCE/DOCKET NUMBER: BIOG-20240/8129-040
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 854-3660
/ TELEFAX: (415) 854-3694
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1442
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA
/ DESCRIPTION: DNA
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM:
/ IMMEDIATE SOURCE:
/ CLONE:
/ FEATURE:
/
/ US-08-152-483B-8
/
/ Query Match 3.9%; Score 35.4; DB 1; Length 1442;
/ Best Local Similarity 51.8%; Pred. No. 0.77; Mismatches 0; Indels 0; Gaps 0;
/ Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/
/ QY 74 CATCACTTCGTGTAGTTACAAACGCTTATCATGGAGCGCTGATCTGGTACAAAGCTCAAG 133
/ Db 542 CACCGGCTCGTAGTGTGCTGACGACGCTGATCCACGAGCGCTGGCGGTTGGGCTACATG 483
/
/ QY 134 CCCCTCCTTTGAATCAGCGGCCACCGGTTACTGTGTGAACTCGCCGCTCCGGGATC 193
/ Db 482 CAGTTCGAGCGCGTCACGAAACACGCGGTAAACCGCGCTGTATGTCGCGCGCGGTGACC 423
/
/ QY 194 GACCCAGCAATCAGCGCGCTTGAACCGTGCAGC 230
/ Db 422 GTCACGCGCGTGACCGCGCGGTGGTGGCCCGCGCG 386
/
/ RESULT 10
/ US-08-152-483B-6/c
/ Sequence 6, Application US/08152483B
/ Patent No. 5529909
/ GENERAL INFORMATION:
/ APPLICANT: Della-Cioppa, Guy
/ APPLICANT: Kumagai, Monto
/ TITLE OF INVENTION: TYROSINASE-ACTIVATOR
/ TITLE OF INVENTION: PROTEIN FUSION ENZYME
/ NUMBER OF SEQUENCES: 9
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Pennie & Edmonds
/
/
/ STREET: 2730 Sand Hill Road
/ CITY: Menlo Park
/ STATE: California
/ COUNTRY: U.S.A.
/ ZIP: 94025
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent in Release #1.0,
/ SOFTWARE: Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/152,483B
/ FILING DATE: No. 5529909ember 12, 1993
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 7/857,602
/ FILING DATE: March 30, 1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 923,692
/ FILING DATE: July 31, 1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 600,244
/ FILING DATE: October 22, 1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 641,617
/ FILING DATE: January 16, 1991
/ NAME: Halluin, Albert P.
/ REGISTRATION NUMBER: 25,227
/ REFERENCE/DOCKET NUMBER: BIOG-20240/8129-040
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 854-3660
/ TELEFAX: (415) 854-3694
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4009
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: circular
/ MOLECULE TYPE: DNA
/ DESCRIPTION: NO
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM:
/ IMMEDIATE SOURCE:
/ CLONE:
/ FEATURE:
/
/ US-08-152-483B-6
/
/ Query Match 3.9%; Score 35.4; DB 1; Length 4009;
/ Best Local Similarity 51.6%; Pred. No. 1.6; Mismatches 0; Indels 0; Gaps 0;
/ Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/
/ QY 74 CATCACTTCGTGTAGTTACAAACGCTTATCATGGAGCGCTGATCTGGTACAAAGCTCAAG 133
/ Db 2780 CACCGGCTCGTAGTGTGCTGACGACGCTGATCCACGAGCGCTGGCGTTGCGCATCACATG 2721
/
/ QY 134 CCCCTCCTTTGAATCAGCGGCCACCGGTTACTGTGTGAACTCGCCGCTCCGGGATC 193
/ Db 2720 CAGTTCGAGCGCGTCCACGAAACACGCGCGTAAACCGCGCGTGTATGTCGCGCGGTGACC 2661
/
/ QY 194 GACCCAGCAATCAGCGCGCTTGAACCGTGCAGC 230
/ Db 2660 GTCACGCGCGTGACCGCGCGTGGTGGCCCGCGCG 2624
/
/ RESULT 11
```

US-08-152-483B-2/c  
; Sequence 2, Application US/08152483B  
; Patent No. 5529909  
; GENERAL INFORMATION:  
; APPLICANT: Della-Cioppa, Guy  
; APPLICANT: Kumagai, Monto  
; TITLE OF INVENTION: TYROSINASE-ACTIVATOR  
; TITLE OF INVENTION: PROTEIN FUSION ENZYME  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 2730 Sand Hill Road  
; CITY: Menlo Park  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0,  
; SOFTWARE: Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/152,483B  
; FILING DATE: No. 5529909ember 12, 1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 7/857,602  
; FILING DATE: March 30, 1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 923,692  
; FILING DATE: July 31, 1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 600,244  
; FILING DATE: October 22, 1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 641,617  
; FILING DATE: January 16, 1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 737,899  
; FILING DATE: July 26, 1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Halluin, Albert P.  
; REGISTRATION NUMBER: 25,227  
; REFERENCE/DOCKET NUMBER: BIOG-20240/8129-040  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 854-3660  
; TELEFAX: (415) 854-3694  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4294  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; DESCRIPTION: DNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM:  
; IMMEDIATE SOURCE:  
; CLONE:  
; FEATURE:  
; US-08-152-483B-2  
Query Match 3.9%; Score 35.4; DB 1; Length 4294;  
Best Local Similarity 51.6%; Pred. No. 1.7;  
Matches 81; Conservative 0; Mismatches 76; Indels 0; Gaps 0;  
Qy 74 CATCACTTCGTGTAGTTATCACACGCTTATCATGGAGCCTGGATCTGGTACAGCTCAAG 133  
Db 3590 CACCGGCTCGTAGTGGCTGACGACGCTGATCCACGAGCGCTGGCGGTCATCACATG 3531

Qy 134 CCCCTCCTTGAATCAGCGCGGCCACGCGGTTACTCTGTGCAACTCGCGGCTCCGGGATC 193  
Db 3530 CAGTTTCGACGCGCTCCACGAAACACGCGGTAAACCCGCGCGTATGGCGCGCCGCTGACC 3471  
Qy 194 GACCCACGACCAATCCAGGCGGTTGAAACCGTCGACG 230  
Db 3470 GTCACCGCGGTGACCGCGCGCGTGTGGCCCCCGCCG 3434  
RESULT 12  
US-09-949-016-16445  
; Sequence 16445, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16445  
; LENGTH: 42094  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1) --(42094)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-16445  
Query Match 3.9%; Score 35.4; DB 4; Length 42094;  
Best Local Similarity 50.9%; Pred. No. 8.6;  
Matches 84; Conservative 0; Mismatches 81; Indels 0; Gaps 0;  
Qy 398 CACGTGCTTCTCAGCTTCGACAAAGTATATGAGATGCTCGGAGTTTGGGAGATTGT 457  
Db 10870 CATGTAGCTTTTAAAGTTTAGTTAAATTAAGTAAATAATTAAGCCTAGGCAACATAGC 10929  
Qy 458 GAGTTTTCATCTCATGAACAAGAAATGGGACGATGAGTTTATTGAAGATGGACCAAAA 517  
Db 10930 AAGACCTCATCTCTGAAAAAAGAAAAAGATAAAGATTAGCTGACATGGTGTGTACA 10989  
Qy 518 TTCATGAAGGACGCTTTTACCAAAATTTGCCATAGAGGATTAC 562  
Db 10990 TGCCTGTAGTCCAGCTACTCAGGAGGCTGAGATGGGAGGATTAC 11034  
RESULT 13  
US-09-621-976-2813/c  
; Sequence 2813, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 2813  
; LENGTH: 832  
; TYPE: DNA



**This Page Blank (uspro)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: September 12, 2005, 04:46:01 ; Search time 2020.05 Seconds  
(without alignments)  
2964.938 Million cell updates/sec

Title: US-10-780-002-36  
Perfect score: 912  
Sequence: 1 atcagcgtattctcaacaa.....ttttcgggcaactttc 912

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7351250 seqs, 3283620254 residues

Total number of hits satisfying chosen parameters: 14702500

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US05\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:\*
- 22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq:\*
- 24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description         |
|------------|-------|-------|--------|----|---------------------|
| 1          | 912   | 100.0 | 912    | 21 | US-10-780-002-36    |
| 2          | 481.4 | 52.8  | 483    | 10 | US-09-770-961-854   |
| 3          | 243   | 26.6  | 989    | 21 | US-10-780-002-32    |
| 4          | 241.6 | 26.5  | 792    | 9  | US-09-938-842A-1320 |
| 5          | 241.6 | 26.5  | 792    | 11 | US-09-938-842A-1320 |
| 6          | 232.6 | 25.5  | 1079   | 21 | US-10-780-002-1     |
| 7          | 223.4 | 24.5  | 933    | 18 | US-10-424-599-64748 |

|    |       |      |      |    |                      |                    |
|----|-------|------|------|----|----------------------|--------------------|
| 8  | 222.6 | 24.4 | 837  | 21 | US-10-780-002-34     | Sequence 34, Appl  |
| 9  | 218.8 | 24.0 | 792  | 9  | US-09-938-842A-217   | Sequence 217, App  |
| 10 | 218.8 | 24.0 | 792  | 11 | US-09-938-842A-217   | Sequence 217, App  |
| 11 | 214.2 | 23.5 | 792  | 9  | US-09-938-842A-1551  | Sequence 1551, Ap  |
| 12 | 214.2 | 23.5 | 792  | 11 | US-09-938-842A-1551  | Sequence 1551, Ap  |
| 13 | 214.2 | 23.5 | 870  | 21 | US-10-780-002-33     | Sequence 33, Appl  |
| 14 | 205.4 | 22.5 | 777  | 21 | US-10-738-922-1      | Sequence 1, Appli  |
| 15 | 205.4 | 22.5 | 777  | 21 | US-10-738-927-1      | Sequence 1, Appli  |
| 16 | 204.8 | 22.5 | 1101 | 19 | US-10-437-963-62048  | Sequence 62048, A  |
| 17 | 202.6 | 22.2 | 1158 | 18 | US-10-424-599-85023  | Sequence 85023, A  |
| 18 | 199.6 | 21.9 | 792  | 21 | US-10-780-002-35     | Sequence 35, Appl  |
| 19 | 198.6 | 21.8 | 777  | 21 | US-10-738-922-2      | Sequence 2, Appli  |
| 20 | 198.6 | 21.8 | 777  | 21 | US-10-738-927-2      | Sequence 2, Appli  |
| 21 | 194.4 | 21.3 | 771  | 9  | US-09-938-842A-1278  | Sequence 1278, Ap  |
| 22 | 194.4 | 21.3 | 771  | 11 | US-09-938-842A-1278  | Sequence 1278, Ap  |
| 23 | 194.4 | 21.3 | 771  | 21 | US-10-780-002-40     | Sequence 40, Appl  |
| 24 | 192.2 | 21.1 | 834  | 21 | US-10-780-002-39     | Sequence 39, Appl  |
| 25 | 191.6 | 21.0 | 1035 | 18 | US-10-425-114-12800  | Sequence 12800, A  |
| 26 | 185.2 | 20.3 | 783  | 21 | US-10-780-002-38     | Sequence 38, Appl  |
| 27 | 176.8 | 19.4 | 819  | 21 | US-10-780-002-37     | Sequence 37, Appl  |
| 28 | 172.6 | 18.9 | 901  | 21 | US-10-780-002-41     | Sequence 41, Appl  |
| 29 | 171.2 | 18.8 | 1069 | 19 | US-10-787-701-9317   | Sequence 9317, Ap  |
| 30 | 149.6 | 16.4 | 996  | 19 | US-10-767-701-15339  | Sequence 15339, A  |
| 31 | 146.4 | 16.1 | 653  | 18 | US-10-424-599-52897  | Sequence 52897, A  |
| 32 | 145.2 | 15.9 | 1148 | 19 | US-10-437-963-29636  | Sequence 29636, A  |
| 33 | 141.2 | 15.5 | 786  | 19 | US-10-437-963-62697  | Sequence 62697, A  |
| 34 | 134.8 | 14.8 | 1173 | 19 | US-10-437-963-89948  | Sequence 89948, A  |
| 35 | 132   | 14.5 | 1007 | 18 | US-10-425-114-21611  | Sequence 21611, A  |
| 36 | 132   | 14.5 | 1018 | 20 | US-10-425-115-119846 | Sequence 119846, A |
| 37 | 128.8 | 14.1 | 1025 | 18 | US-10-425-114-13774  | Sequence 13774, A  |
| 38 | 128.8 | 14.1 | 1092 | 20 | US-10-425-115-70617  | Sequence 70617, A  |
| 39 | 126   | 13.8 | 818  | 18 | US-10-425-114-16669  | Sequence 16669, A  |
| 40 | 126   | 13.8 | 830  | 20 | US-10-425-115-156630 | Sequence 156630, A |
| 41 | 118.4 | 13.0 | 1233 | 19 | US-10-437-963-19868  | Sequence 19868, A  |
| 42 | 118.2 | 13.0 | 920  | 18 | US-10-424-599-128189 | Sequence 128189, A |
| 43 | 117.2 | 12.9 | 439  | 18 | US-10-424-599-81721  | Sequence 81721, A  |
| 44 | 117   | 12.8 | 901  | 20 | US-10-739-930-3765   | Sequence 3765, Ap  |
| 45 | 116   | 12.7 | 907  | 18 | US-10-425-114-24341  | Sequence 24341, A  |

ALIGNMENTS

RESULT 1  
US-10-780-002-36  
; Sequence 36, Application US/10780002  
; Publication No. US20050034196A1  
; GENERAL INFORMATION:  
; APPLICANT: Klesbig, Daniel F.  
; APPLICANT: Kumar, Dhirenda  
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN  
; FILE REFERENCE: 3670-PO2652W0  
; CURRENT APPLICATION NUMBER: US/10/780,002  
; CURRENT FILING DATE: 2004-02-17  
; PRIOR APPLICATION NUMBER: PCT/US02/26312  
; PRIOR FILING DATE: 2002-08-16  
; PRIOR APPLICATION NUMBER: 60/312,863  
; PRIOR FILING DATE: 2001-08-16  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 36  
; LENGTH: 912  
; TYPE: DNA  
; ORGANISM: Arabidopsis Thaliana  
US-10-780-002-36

Query Match 100.0%; Score 912; DB 21; Length 912;  
Best Local Similarity 100.0%; Pred. No. 4.7e-278; Indels 0; Gaps 0;  
Matches 912; Conservative 0; Mismatches 0  
QY 1 ATCAGCGTATTCTCAACAAAGCAAGAACCTTAAACCAAAAGAAACGCTACGGAGACA 60  
|||||



Publication No. US20050034196A1  
GENERAL INFORMATION:  
APPLICANT: Klessig, Daniel P.  
APPLICANT: Kumar, Dhirenda  
TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN  
FILE REFERENCE: 3670-PO2652WO  
CURRENT APPLICATION NUMBER: US/10/780,002  
CURRENT FILING DATE: 2004-02-17  
PRIOR APPLICATION NUMBER: PCT/US02/26312  
PRIOR FILING DATE: 2002-08-16  
PRIOR APPLICATION NUMBER: 60/312,863  
PRIOR FILING DATE: 2001-08-16  
NUMBER OF SEQ ID NOS: 58  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 32  
LENGTH: 989  
TYPE: DNA  
ORGANISM: Arabidopsis Thaliana  
US-10-780-002-32

Query Match 26.6%; Score 243; DB 21; Length 989;  
Best Local Similarity 59.6%; Pred. No. 9.7e-66;  
Matches 468; Conservative 0; Mismatches 305; Indels 12; Gaps 3;  
QY 66 AGAGGAAACATCACTTCGTGTAGTTTCAACAACGCTTATCATGAGCGCTGGATCTGGTACA 125  
DB 120 AGAGGAAACAACTTTGTACTAGTACATGGTTCGTCATGCGCGTGGTGTGGTACA 179  
QY 126 AGCTCAAGCCCTCTCTTGAATCAGCGCGGCACCGCGTTACTGTGTCGAACCTCGCGCCT 185  
DB 180 AGGTTAAGCCGCTGTAGAGCGGTGGCCACCGCGTAACTGCTGTGGACTTAGCTGCT 239  
QY 186 CCGGGATCGACCA----CGACCAATCCAGCGCTTGAACCGTCGACGAATCTCCAAAC 242  
DB 240 CCGGAATAGACACAACAGGAGTCGATCACTGACATCCCACTCGCAACAATACTCGGAGC 299  
QY 243 CGTTGATCGAAACCTCAAACTCTTCCAGAGAACGAAGGTAATCTGCTTGGATTCA 302  
DB 300 CATTTGACGAAGCTCTTGACTCTATTGCCAAATGATGAAGGTTGTGCTCGTGTGGACA 359  
QY 303 GCTTCGAGGATCAACATCGCTCTCGCGCGGCACATATTTCCGCGGAAGATTAAAGTTTC 362  
DB 360 GCCTTGGTGGCTTGAACCTTAGCATAGCCATGGAAGTTTCCGGAATAATCTCTGTGC 419  
QY 363 TTGTGTTCTCAACGCTTCTTGGCCGACACAACCGCTGCTTCTCAAGTTCTGGACA 422  
DB 420 CTGTATTCTTGAAGTCTTTCATGCGCGACACCGAACACTCAACATCTCTCGTCTGGACA 479  
QY 423 AGTAT-----ATGGAGATGCTCGAGGTTTGGAGATGTGAGTTTCTCATGTGAA 476  
DB 480 AGTTTGGAAACAACTGCTCAAGAAGATGATGGGACCGAAATTCGAACCTTATGTT 539  
QY 477 CAAGAAATGGAGATGAGTTTATGAAGATGGACCAAAATTCATGAAGGACGCTCTTT 536  
DB 540 CAGACAATTCGGACTGAGTATGTT---TTTGTAGCCTGACTTCATGAAGTTGGTCTCT 596  
QY 537 ACCAAATTTGTCATAGAGGATTAAGAGCTGGCAAAATGTTGCATAGGCAAGGTCAT 596  
DB 597 ACCAGCTTCTCCAGTTGAGGATCTTGAACCTGGATTACTTTTAAAGAGCCAGGATCCT 656  
QY 597 TTTTTCAGAGGATCTCAAGAAAGAAAGTTTGTAGCGAGGAGGATGTTGGTGC 656  
DB 657 TATTTAATACGATTTATCGAAGATGAATAAATCTTCGGATGAAGGATGTTGTTCT 716  
QY 657 AAGCGTTTACGTAATGAGTAGGAAGACAAGAGCCATCCCTGCGATTTCAATTCGTTGGA 716  
DB 717 CTCGAGTTTTCATGTTGTGAAGAGGACAAGGAATTCAGAGAACCGCCAGAGATGA 776  
QY 717 TGATTGATAATTTCAACGCTCGAAAGTCTACAGAGTCGATGCGGAGATCAATGGTGA 776  
DB 777 TGATTGATAATTTTCCGGTGAATTTAGTGTGATGAGATGGAGGAGACAGATCATATGCCAA 836

QY 777 TGCTCTCAAAACCCCAAAACTTTTGACTCTCTCTCTGCTATTGCCACCGATTATGT 836  
DB 837 TGTCTGCAAGCCTCAGCACTCAGTGATTACTTCTGAAATCGGACAAATTCGTTT 896  
QY 837 AATAA 841  
DB 897 AATCA 901  
RESULT 4  
US-09-938-842A-1320  
Sequence 1320, Application US/09938842A  
Patent No. US20020160378A1  
GENERAL INFORMATION:  
APPLICANT: Harper, Jeff  
APPLICANT: Kreps, Joel  
APPLICANT: Wang, Xun  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
TITLE OF INVENTION: SAME, AND METHODS OF USE  
FILE REFERENCE: SRIPI300-3  
CURRENT APPLICATION NUMBER: US/09/938,842A  
CURRENT FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: US 60/227,866  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: US 60/264,647  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/300,111  
PRIOR FILING DATE: 2001-06-22  
NUMBER OF SEQ ID NOS: 5379  
SEQ ID NO 1320  
LENGTH: 792  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1320

Query Match 26.5%; Score 241.6; DB 9; Length 792;  
Best Local Similarity 59.6%; Pred. No. 2.4e-65;  
Matches 466; Conservative 0; Mismatches 304; Indels 12; Gaps 3;  
QY 66 AGAGGAAACATCACTTCGTGTAGTTTCAACAACGCTTATCATGAGCGCTGGATCTGGTACA 125  
DB 14 AGAGGAAACAACTTTGTACTAGTACATGGTTCGTGCGCATGCGCGTGGTGTGATA 73  
QY 126 AGCTCAAGCCCTCTCTTGAATCAGCGCGGCACCGGTTACTGCTCGCAACTCGCGCCT 185  
DB 74 AGGTTAAGCCGCTGTAGAGCGGTGGGCCACCGGTAACCTGCTGTGACTTAGTGTCT 133  
QY 186 CCGGGATCGACCCA---CGACCAATCCAGGCGGTTGAAACCGTCGACGAATACTCCAAAC 242  
DB 134 CCGGATAGACACAACAGAGGTGATCACTGATCCCACTGCGAACATACTCGGAGC 193  
QY 243 CGTTGATCGAAACCTTCAAACTCTTCCAGAGAACGAAGGTAATCTGTTGGATTCA 302  
DB 194 CATTGACGAAGCTCTGACCTCATTTGCCAAATGATGAAGAGTTGTGCTCGTTGGTACA 253  
QY 303 GCTTCGAGGATCAACATCGCTCTCGCGCGGCACATATTTCCGCGGAGATTAAAGTTTC 362  
DB 254 GCTTTGTGGCTTGAACCTTAGCCATAGCCATGGAAGTTTCCCGAAAAAATCTCTGTGC 313  
QY 363 TTGTGTTCTCAAGCCCTTCTTGGCCGACAAACCGTCGCTTCTCACGTTCTGGACA 422  
DB 314 CTGTATTCTGACTGCTTTCATGCGGACCGAACACTCAACATCTTCTGCTTGGACA 373  
QY 423 AGTAT-----ATGGAGATGCTCGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAA 476  
DB 374 AGTTTGAAGCAACATGCTCAAGAAGCATGATGGGACCCGAAATTCGAACCTTATGTT 433  
QY 477 CAAGAAATGGAGATGAGTTTATGAAGATGGGACCAAAATTCATGAAGGACGCTCTTT 536  
DB 434 CAGACAATTCGGAGCTGAGTATGTT---TTTGTAGCCCTGACTTCAATGAAGTTGGTCTCT 490  
QY 537 ACCAAATTTGTCCTCATAGAGGATTACGAGCTGGGCAAAATGTTGCTAGTAGGAGGTCAT 596

|     |  |     |
|-----|--|-----|
| 491 | ACCAGCTTTCTCAGTTGAGGATCTTGAACTGGGATTACTTTTAATGAGGCCAGGATCGT    | 550 |
| 597 | TTTTCACAGGATCTATCAAAGAAAGAAAAGTTTAGCGAGGAAGGATATGGTTTCGGTGC    | 656 |
| 551 | TATTTATTAAAGATTTATCGAAGATGAAAAAATCTCTCGGATGAAGGATATGGGTCTGTTTC | 610 |
| 657 | AACGAGTTTTACGTAATGAGTAGTGAAGACAAGAGCCATCCCTGCGATTTCATTCGTTGGA  | 716 |
| 611 | CTCGAGTTTTCATAGTGTGTAAGAGGACAAGCAATTCAGAAAGAACGCCAGAGATGGA     | 670 |
| 717 | TGATTCGATAATTTCAACGCTCTCGAAAGCTCAGGAGATCGATGGCGGAGATCAATGGTGA  | 776 |
| 671 | TGATTGATAAATTTCCCGGTGAATTTAGTAGTGGAGATGGAGGACAGATCATATGCCAA    | 730 |
| 777 | TGCTCTCCAAACCCCAAAACCTCTTTTGACTCTCTCTGCTATTGTCACCGATTATATGT    | 836 |
| 731 | TGTTCTGCAAGCTCAGCAACTCAGTGATTACTCTCGAAAATCGCGGACAAATTCGTTT     | 790 |
| 837 | AA 838   |     |
| 791 | AA 792   |     |

## RESULT 5

US-09-938-842A-1320  
 ; Sequence 1320, Application US/09938842A  
 ; Publication No. US20040009476A9  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harper, Jeff  
 ; APPLICANT: Kreps, Joel  
 ; APPLICANT: Wang, Xun  
 ; APPLICANT: Zhu, Tong  
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE  
 ; FILE REFERENCE: SCRI1300-3  
 ; CURRENT APPLICATION NUMBER: US/09/938,842A  
 ; CURRENT FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/227,866  
 ; PRIOR FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/264,647  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/300,111  
 ; PRIOR FILING DATE: 2001-06-22  
 ; NUMBER OF SEQ ID NOS: 5379  
 ; SEQ ID NO 1320  
 ; LENGTH: 792  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 US-09-938-842A-1320

|                       |        |  |          |            |        |        |     |      |   |
|-----------------------|--------|--|----------|------------|--------|--------|-----|------|---|
| Query Match           | 26.5%; | Score  | 241.6;   | DB 11;     | Length | 792;   |     |      |   |
| Best Local Similarity | 59.6%; | Prod. No.  | 2.4e-65; |            |        |        |     |      |   |
| Matches               | 466;   | Conservative   | 0;       | Mismatches | 304;   | Indels | 12; | Gaps | 3 |
| Qy                    | 66     | AGGGAACATCACTTCCTGCTAGTTTACAAACGCTTATCATGGAGCCTGGATCTCGTACA      | 125      |            |        |        |     |      |   |
| Db                    | 14     | AGAGAAACAATTTTGTACTGATCATGGTTCTGCCATGGCCGCTGGTCTCGTACA           | 73       |            |        |        |     |      |   |
| Qy                    | 126    | AGCTCAAGCCCTCTCTTGAATCAGCGCGGCACCGCGTTTACTGCTCTCGAACTCGCGCGCT    | 185      |            |        |        |     |      |   |
| Db                    | 74     | AGGTTAAGCCGCTGTAGAGGGGTGGGCACCGCGTAACCTGCTGGACTAGTGTGCT          | 133      |            |        |        |     |      |   |
| Qy                    | 186    | CCGGGATCGACCCA---CGACAATTCAGGCGGTTTGAACCGTTCGACGAATACTCCAAAC     | 242      |            |        |        |     |      |   |
| Db                    | 134    | CCGGAATAGACACACGAGGTCGATCACTGCATCCCCACATCGCAACAATACTCGGAGC       | 193      |            |        |        |     |      |   |
| Qy                    | 243    | CGTTGATCGAACCTCAAATCTCTTCAGAGAACGGAAGAGTAATCTCTGGTTGATTCACA      | 302      |            |        |        |     |      |   |
| Db                    | 194    | CATTGACGAAGTCTCTGACCTCATTTGCCAAATGATGAAAAGGTTGTCTCTGTTGTACA      | 253      |            |        |        |     |      |   |
| Qy                    | 303    | CTTTCCGGAGGCATCAACAATCGCTCTTCGCCGCCGACATATTTCCGGCGAAGATTAAGGTTTC | 362      |            |        |        |     |      |   |

|  |     |  |                              |                      |     |
|--|-----|--|------------------------------|----------------------|-----|
| Db   | 254 | GCTTTGGTGGCTTGA  | CTTAGCCATAGCCATAGGAAAGATTTC  | CCGAAAAAATCTCTGTGC   | 313 |
| Qy   | 363 | TTTGHTTCCTCAACGCCTTCTTGCCGACACAA                               | CCCAACGTCCTTCTCA             | CGTTCGTGGACA         | 422 |
| Db   | 314 | CTGTATTCTTGACTGCTTTTCATGCGGACAC                                | CCGAACATCA                   | CCCATCTCTCGTCTTGGACA | 373 |
| Qy   | 423 | AGTAT-----ATGGAGATGCGTCGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAA    | 476                          |                      |     |
| Db   | 374 | AGTTTGGAAAGCAACATCGCTCAAGAAAGATGGATGGGAC                       | CCGAATTCGAACCTTATGTTT        | 433                  |     |
| Qy   | 477 | CAAGAAATGGGACGATGAGTTTATTTGAAGATGGGACCAAAAATTCATGAAGGCA        | CGTCTTTT                     | 536                  |     |
| Db   | 434 | CAGACAATCCGGACTGAGTATGTT---TTTTAGCCCTGACTTTCATGAAGTTGGGTC      | TCT                          | 490                  |     |
| Qy   | 537 | ACCAAAATGTGCCATAGAGATTACAGAGCTGGCAAAAATGTTTGCATAGGCAAGGTCAT    | 596                          |                      |     |
| Db   | 491 | ACCAGCTTCTCCAGTTTGAGGATCTTGAACCTGGGATTA                        | CTTTTAATGAGGCCAGGATCGT       | 550                  |     |
| Qy   | 597 | TTTTTCACAGAGGATCTATCAAGAAAGAAAAGTTT                            | TAGCGAGGAGGATATGTTCCGTGC     | 656                  |     |
| Db   | 551 | TAITTTATTAAGATTATCGAAGATGAAAACCTTCTCGGATGAAGGATATGGTCTGTC      | 610                          |                      |     |
| Qy   | 657 | AACGAGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCCTGCGAATTCATTCGTTGA    | 716                          |                      |     |
| Db   | 611 | CTCGAGTTTTCATAGTGTGTAAAGAGGACAAAGCAA                           | TTCCAGAGAAGCCAGAGATGGA       | 670                  |     |
| Qy   | 717 | TGATTGATTAATTCAAAGCTCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTGA  | 776                          |                      |     |
| Db   | 671 | TGATTGATAATTTTTCCGGTGAATTTAGTGATGGGAGATGGAGGAGACAGATCATATGCCAA | 730                          |                      |     |
| Qy   | 777 | TGCTCTCCAAACCCCAAAAACCTCTTTGACTCTCTCTGCTATTGCCCACCGATTATATGT   | 836                          |                      |     |
| Db   | 731 | TGTTCTGCAAGCCTCAGCAACTCAGTGATTA                                | CTTCTCGAAATTCGGGACAAATTCGTTT | 790                  |     |
| Qy   | 837 | AA   | 838                          |                      |     |
| Db   | 791 | AA   | 792                          |                      |     |
| RESULT 6   |     |  |                              |                      |     |
| US-10-780-002-1  |     |  |                              |                      |     |
| ; Sequence 1, Application US/10780002                      |     |  |                              |                      |     |
| ; Publication No. US20050034196A1                          |     |  |                              |                      |     |
| ; GENERAL INFORMATION:                                     |     |  |                              |                      |     |
| ; APPLICANT: Klessig, Daniel F.                            |     |  |                              |                      |     |
| ; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN |     |  |                              |                      |     |
| ; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN |     |  |                              |                      |     |
| ; FILE REFERENCE: 3670-PO2652WO                            |     |  |                              |                      |     |
| ; CURRENT APPLICATION NUMBER: US/10/780,002                |     |  |                              |                      |     |
| ; CURRENT FILING DATE: 2004-02-17                          |     |  |                              |                      |     |
| ; PRIOR APPLICATION NUMBER: PCT/US02/26312                 |     |  |                              |                      |     |
| ; PRIOR FILING DATE: 2002-08-16                            |     |  |                              |                      |     |
| ; PRIOR APPLICATION NUMBER: 60/312,863                     |     |  |                              |                      |     |
| ; PRIOR FILING DATE: 2001-08-16                            |     |  |                              |                      |     |
| ; NUMBER OF SEQ ID NOS: 58                                 |     |  |                              |                      |     |
| ; SOFTWARE: FastSeq for Windows Version 3.0                |     |  |                              |                      |     |
| ; SEQ ID NO 1  |     |  |                              |                      |     |
| ; LENGTH: 1079   |     |  |                              |                      |     |
| ; TYPE: DNA  |     |  |                              |                      |     |
| ; ORGANISM: Nicotiana tabacum                              |     |  |                              |                      |     |
| US-10-780-002-1  |     |  |                              |                      |     |

RESULT 6

```

US-10-780-002-1
; Sequence 1, Application US/10780002
; Publication No. US2005034196A1
; GENERAL INFORMATION:
; APPLICANT: Klessig, Daniel F.
; APPLICANT: Kumar, Dhirenda
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN
; TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, SAbF2, AND METHODS OF USE THEREOF
; FILE REFERENCE: 3670-PO2652WO
; CURRENT APPLICATION NUMBER: US/10/780,002
; PRIORITY FILING DATE: 2004-02-17
; PRIORITY APPLICATION NUMBER: PCT/US02/26312
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/312,863
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-10-780-002-1

```

|    | Query Match           | 25.5%   | Score 232.6     | DB 21    | Length 1079 |
|----|-----------------------|---|-----------------|----------|-------------|
|    | Best Local Similarity | 56.3%   | Pred. No. 2e-62 |          |             |
|    | Matches 458           | Conservative 0  | Mismatches 349  | Indels 6 | Gaps 17     |
| Qy | 33                    | AAACAAAAGAAAACGCTACCGGAGACATATGGAGAGAAACATCACTTCGTGTTAGTTC    | 92              |          |             |
| Dd | 13                    | AAAAGAACTTCAACAGGCATAAATTCAAATGAAGGAAGGAAACACTTGTGTTTAGTAC    | 72              |          |             |
| Qy | 93                    | ACAACGCTTATCATGGAGCCTGGATCTGGTATCAAGCTCAAGCCCGCTCTTGAATCAGCCG | 152             |          |             |



Db 73 ATGTCATGCCATGAGGTTGGAGTTGGTACAAAGCTAAAGCCACTGCTAGAAAGTCGAC 132  
QY 153 GCACCGCGTTACTGCTCGAAGTCCGCGCTCCCGGATCGACCCACGACCAATCCAGG 212  
Db 133 GCCATRAAGTTACAGCCCTTGATTTAGCAGCTTCTGGCACTGATTTGAGAAAATAGAGG 192  
QY 213 CGGTTGAAACCGTCGAGCAATACCTCAAAACCGTTGATCGAAACCCCTCAAAATCTTTCCAG 272  
Db 193 AGCTTCGCACACTTTATGATGATTAACCTTTGCAATGATGAGTTGATGGAATCTTTTCAG 252  
QY 273 AGAACCAAGAGGTAATCTCGTTGGATTGAGTTCCGAGGCAATCAACATCGCTCTCGCG 332  
Db 253 CAGATGAGAAGGTTATATTAGTGGGCATAGTCTTGGTGTATGAATTTGGGACTTGTCTA 312  
QY 333 CCACATATTTCCGGCGAAGATTAAAGTTCTTGTTCTCTCAACGCTTCTTGCCCGACA 392  
Db 313 TGGAAAGTATCCAAAGAGTCTATGCTGCTGTTTCTTGCTGCTTTTCATGCTGATTT 372  
QY 393 CAACCCAGTGCCTTCTCAGCTTCTGCAAGATGATGAGATGCTCGAG-----GPT 446  
Db 373 CTGTTCAACAACCTCTCTGTTGTTTGGAAACAGTATAATGAGGGGACGCCAGCGGAAT 432  
QY 447 TGGAGATTGTGATGTTTCTCATCTCAAGAAACAGAAATGGGAGGATGAGTTTATTGAAGA 506  
Db 433 GGTGGATATCTCAGTTTTCACATATGTTCCCTGAAGAGCCACTGACATCCATGTTTT 492  
QY 507 TGGGACCAAAATTCATGAAGGACGCTTTTACCAGAAATTTGCCATAGAGGATTACGAGC 566  
Db 493 TTGGCCAAAGTTCTTGGCTCAAGCTCTACAGCTATGCTCTCTGAGGATCTTGCAAT 552  
QY 567 TGGCAAAATGTTGTCATAGGCAAGGTCATTTTTCACAGAGGATCTATCAAAAGAAAGAA 626  
Db 553 TAGCATCATATTTGGTGAGACCAAGCTCTTTGTTATGGAAGACCTATCGAAGGCCAAGT 612  
QY 627 AGTTTACGAGGAGGATGATGTTGCTGCAAGAGTTTACGTAAGTACGATGATGAGACA 686  
Db 613 ATTTCACAGATGAACGTTTGGATCAGTGAAGAGGTTTACATTTGTGCTACCTGAGGATA 672  
QY 687 AAGCCATCCCTCGGATTTTCAATCGTTGATGATGATTAATTTCAACGCTCGAAGTCT 746  
Db 673 AAGGCATACAGAAAGTTTCCAGCGATGGCAATTTGCAACATTTGGTGTCACTGAAGCAA 732  
QY 747 ACGAGATCGATGGCGGAGATCAGATGTTGATGCTCTCCAAACCCCAAAACCTCTTTGACT 806  
Db 733 TAGAGATTAAGGTCGTCGATCAGTCGCAATGCTATGCGAGCCCAAAACCTTTGCGCT 792  
QY 807 CTCTCTCTGATTTGCCACCGATTATATGTAAT 839  
Db 793 CTCTCTTGGAAATTTGCCATAAATACAACTGAT 825

RESULT 7

US-10-424-599-64748  
; Sequence 64748, Application US/10424599  
; Publication No. US20040031072A1

GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; NUMBER OF FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 64748

; LENGTH: 933

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_29480C.1

US-10-424-599-64748

Query Match 24.5%; Score 223.4; DB 18; Length 933;  
Best Local Similarity 55.9%; Pred. No. 1.6e-59;  
Matches 447; Conservative 0; Mismatches 346; Indels 6; Gaps 1;

QY 75 ATCACTTCTGTTAGTTTACAAAGCTTATCATCGAGCCCTGGATCTGGTACAAAGCTCAAGC 134  
Db 44 AGCATTTGTTTCTGTTGATGCGGATGCCATGGAGCATGGTGTGGTATTAAGCTCAAGC 103  
QY 135 CCTCTTTGAAATACGCGCCACCGCTTACTGCTGTGAACTCCGCGCTCCGGGATCG 194  
Db 104 CAGGTTTGAATCTGCTGGGCAAGGTCACAGCTCTGACCTTGAGCTTCTGAGGCCA 163  
QY 195 ACCCAACCAATCCAGCCGTTGAAACCGTCGACGAATCTCCAAACCGTTGATCGAAA 254  
Db 164 ACATGAAGAAATTTGAAGATGTTGACATTTTTCACAGTATACTGAGCCTTTGTTGTTTC 223  
QY 255 CCTCAAAATCTTCCAGAGAACGAGAGTAATTTCTGGTTGGATTCAGCTTCGAGGCCA 314  
Db 224 TATTGGACACAATTTCCCTCAAAATGAAAAGGTAGTTTCTAGTTGGTCCACAGCTTTGGAGGC 283  
QY 315 TCAACATCGCTCGCGCCGACATATTTCCGGCGAAGATTAAAGTTCCTTGTTCTCTCA 374  
Db 284 TGAACATAGACTTCCCATGAGAAATTTCCGAAAAGGTAGCAGTTGGTGTGTTTCTTAA 343  
QY 375 AGCCTTCTTGGCCGACACAAACCCACGTCCTTCTCACGTTCTGACAAAGTATATGGAGA 434  
Db 344 CAGCTTTTGTCCAGACGTTGAAACACACCCCATCTTATGCTCTTGGAAAGTACAGCGAGA 403  
QY 435 TGCCTGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAACAAAGAAATGGGACGATGA 494  
Db 404 GGACCCCGTTAGCTGATGCTAGACACTGAATTTGCTCCAAGTGGAAACAAACAA 463  
QY 495 GTTTATTGAAGTGGGACCAAAATTCATGAAGGACGCTTTTACCAGAAATTTGTCCTCATAG 554  
Db 464 TGTCTTTT-----GGCCCCCACTTCTTGTCCGACAAGCTCTACCAACTATCCCCAATGT 517  
QY 555 AGGATTACGAGCTGGCAAAATTTGTCATAGGCAAGGCTCATTTTTCACAGAGGATCTAT 614  
Db 518 AGGATTTTGAATTTGGCCAAAGCTTTTAAGGCCATCATCACTCTTTCATGGAAGACTTGA 577  
QY 615 CAAAGAAAGAAAGTTTACGAGGAGGATATGGTTGCGTGCACGAGTTTACGTAATGA 674  
Db 578 TAAACAAAGAAAGCTTCTCCAAAGAGGATATGGTGCAGTGCAGCTGCTTTATTTGTTT 637  
QY 675 GTAGTGAAGACAAAGCCATCCCTCGGATTTTCATTCGTTGGATGATGATTAATTTCAAGC 734  
Db 638 GCACTGAGGACCTTGCATTTCCATTTGGAATATACAGCTCTTTCATGATCCAAATTTGGGT 697  
QY 735 TCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGATGATGCTCTCCAAACCCCAAA 794  
Db 698 TCAATGAAGTTGTAGAGATCAAAGGCGCAGACCATATGGCTATGCTTTCGAAAGCCCAAG 757  
QY 795 AACTCTTTGACTCTCTCTGCTATGTCACCGATATATGTAATTAATCTTTAAGTCCGTT 854  
Db 758 AACTATTGATTTCCCTCCAGCAGATAGCGACTAAATATGCAATATGCAATTCACGTAAGCATGCT 817  
QY 855 TTACTTTTCTCATCGTT 873  
Db 818 TTAATTAGTTTATTTGTT 836

RESULT 8

US-10-780-002-34

; Sequence 34, Application US/10780002

; Publication No. US20050034196A1

GENERAL INFORMATION:

; APPLICANT: Klesig, Daniel F.

; APPLICANT: Kumar, Dhirenda

; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN

; TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, SAMP2, AND METHODS OF USE THEREOF

; FILE REFERENCE: 3670-PO2652WO

; CURRENT APPLICATION NUMBER: US/10/780,002

```

; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: PCT/US02/26312
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/312,863
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 837
; TYPE: DNA
; ORGANISM: Arabidopsis Thaliana
US-10-780-002-34

```

| Query Match           | 24.4%          | Score 222.6  | DB 21    | Length 837 |
|-----------------------|----------------|--|----------|------------|
| Best Local Similarity | 56.0%          | Pred. No. 2.6e-59  |          |            |
| Matches 464           | Conservative 0 | Mismatches 359   | Indels 6 | Gaps 2     |
| Qy                    | 16             | AACAAGCAAAAGAACCTTTAACCAAAAAGAAAACGCTACGGAGACATATGGAGAGCAAAACA | 75       |            |
| Db                    | 9              | AAGNAGTCGAGAGACAAGTTTAAAGCTAGAGAAGAAATGAGTCAGGAGAGAGAGCA       | 68       |            |
| Qy                    | 76             | TCACCTTCGTGTTAGTTCAACAACGCTTATCATGGAGCCTGGATCTGGTACAAGCTCAAGCC | 135      |            |
| Db                    | 69             | ACACGTGCTTCTAGTACATGCTGTGCCATGGCGCTGTGCTGTGTACAAAGGTTAAAGCC    | 128      |            |
| Qy                    | 136            | CCTCCTTGAATCAGCCGGCCACCGGTTACTGCTGTGAACTCGCGCCTCCGGGATCGA      | 195      |            |
| Db                    | 129            | GCAGCTCGAGGCTTCTGGCACCGGTAACCGCGTAGATCTAGCTGCTCCGGTATAGA       | 188      |            |
| Qy                    | 196            | C---CCAGCACCAATCCAGGCGGTTGAAACCGTCGAGAACTACTCCAAACGGTTGATCGA   | 252      |            |
| Db                    | 189            | CATGACCAGGTCATACAGATATATCCACATCGCAACATCTCAGAGCCATTGATGCA       | 248      |            |
| Qy                    | 253            | AACCTCAATCTCTTCAGAGAACGAAGGTAATTTCTGTTGGATTACGCTTCGAGG         | 312      |            |
| Db                    | 249            | GCTAATGACCTCACTACCAAGATGATGAAGGTTGTGCTGTTGGTCATAGCTTAGGAGG     | 308      |            |
| Qy                    | 313            | CATCAACATCGCTCTCGCGCGGCACATATTTCCGGCGAAGATTAAGGTTCTTGTGTTCC    | 372      |            |
| Db                    | 309            | TTTGAGTTTAGCTATGGCCATGGATATGTTTCCGACCAGAAATCTGTCTTGCTTTGT      | 368      |            |
| Qy                    | 373            | CAACGCTTCTTTCGCGACACAACCCAGTGCGCTTCTCACGTTCGAGCAAGTATATGGA     | 432      |            |
| Db                    | 369            | GACTGCTATGATGCCAGACACCAACAACACTCAACCTCTCGTATGGGTAAGACTAAAGAA   | 428      |            |
| Qy                    | 433            | GATGCTTGGAGGTTTGGAGATTGTGA--GTTTTCATCTCATGAACAAGAAATGGAC       | 489      |            |
| Db                    | 429            | AGAACTTCACAGAGGATGGTTAGACACCGTGTTCAGAGCAGAAACCTGATTTTCC        | 488      |            |
| Qy                    | 490            | GATGAGTTTATTTGAAGATGGACCACAAATTCATGAAGSCACGTCCTTTACCAAAATGTPCC | 549      |            |
| Db                    | 489            | TAGCGAGTTTGGATTTTGGACCAGAAATTCATGCCAAGAACTGTATCAGTTGTCTCC      | 548      |            |
| Qy                    | 550            | CATAGAGGTTACAGCTGGCAAAAATGTTGCAATAGGCAAGGTCATTTTTCACAGAGGA     | 609      |            |
| Db                    | 549            | AGTCCAAGATCTGAAATGGCGAAAATTTGCGTAGGGCAACCCATGATTAAGAAAGA       | 608      |            |
| Qy                    | 610            | TCTATCAAGAAAGAAAAGTTTAGCGAGAAAGGATATGGTTCCGGTGCACACGAGTTTACGT  | 669      |            |
| Db                    | 609            | TATGCCAGAGAGAGAGCTTCAGTGAGGAAGGATACGGATCCGTTACACGTATATTAT      | 668      |            |
| Qy                    | 670            | AATGAGTAGTGAAGACAAAGCCATCCCTCGGATTTTCAATCGTTGGATGATGATTAATTT   | 729      |            |
| Db                    | 669            | TGTTATGGGAAAGGATCTTGTGTCAACCCGAGAGATTACCAAGGATCGATGATCAGCAACT  | 728      |            |
| Qy                    | 730            | CAACGTCTCGAAGGTCATACGAGATCGATGGCGGAGATCCACATGGTGATGCTCTCCAAACC | 789      |            |
| Db                    | 729            | TCCCCAAAAGAGGTAATGGAGATCAAAAGACGAGATCATATGCCAATGTTCTCCAGCC     | 788      |            |
| Qy                    | 790            | CCAAAAACTCTTTGACTCTCTCTCTGCTATTGCCACCGATTATATGTA               | 838      |            |
| Db                    | 789            | TCAACAACTATGTGCTCTTCTCTGGAGATTGCAAAATAAATATGCTAA               | 837      |            |

RESULT 9  
US-09-938-842A-217  
Sequence 217, Application US/09938842A  
Patent No. US20020160378A1  
GENERAL INFORMATION:  
APPLICANT: Harper, Jeff  
APPLICANT: Kteps, Joel  
APPLICANT: Wang, Xun  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
TITLE OF INVENTION: SAME, AND METHODS OF USE  
FILE REFERENCE: SCRIPT300-3  
CURRENT APPLICATION NUMBER: US/09/938,842A  
CURRENT FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: US 60/227,866  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: US 60/264,647  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/300,111  
PRIOR FILING DATE: 2001-06-22  
NUMBER OF SEQ ID NOS: 5379  
SEQ ID NO 217  
LENGTH: 792  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-938-842A-217

| Query Match           | 24.0%;          | Score 218.8;  | DB 9;     | Length 792; |
|-----------------------|-----------------|---|-----------|-------------|
| Best local Similarity | 56.6%;          | Pred. No. 4.1e-58;  |           |             |
| Matches 447;          | Conservative 0; | Mismatches 337;   | Indels 6; | Gaps 2      |
| Qy                    | 55              | GAGACATATGGAGAGAAACATCACTTCGTGTGTAGTTTCAACAACGCTTATCATGTGAGCCTG | 114       |             |
| Db                    |                 |   |           |             |
| Db                    | 3               | GAGTGAGGAGAGAGAAACACACGTCTCTAGTACATCGTGTCTTCCATGGCGCTG          | 62        |             |
| Qy                    | 115             | GATCTGTTCAAGCTCAAGCCCTCTCTTGTAATCAGCCGCGCACCGCGTTTACTGCTGTGCA   | 174       |             |
| Db                    |                 |   |           |             |
| Db                    | 63              | GTGCTGTTCAAGGTTAAGCGCGACAGCTCGAGGCTTCTGGCCACGCGTAACCGCGGTAGA    | 122       |             |
| Qy                    | 175             | ACTCGCGCCTCCGGGATCGAC---CACAGCAACATCCAGCGCTTGAAACCGTCGCAGCA     | 231       |             |
| Db                    | 123             | TCTAGCTGCTCCGGTATAGACATGACAGGTCAATCACGATATATCCAATGCGAACA        | 182       |             |
| Qy                    | 232             | ATACTCCAAACCGTTGATCGAAACCTCAAATCTCTTCCAGAGAACGAAGAGTTAATCT      | 291       |             |
| Db                    | 183             | ATACTCAGAGCCTATGATGACAGCTTAATGACCTCACTACAGATGATGAGAAAGTTGTGCT   | 242       |             |
| Qy                    | 292             | GGTTGGATTGAGCTTCGGAGGCATCAACATCGCTCTCGCGCCGACATATTTCCGCGCAA     | 351       |             |
| Db                    | 243             | TGTTGGTTCATAGCTTTAGGAGGTTTGAGTTTAGCTATGCGCCATGGATATGTTTTCGACCAA | 302       |             |
| Qy                    | 352             | GATTAAAGTTCTTGTTTCTTCAAACGCTCTTCCGCCGACACAAACCCAGTGCCTTCTCA     | 411       |             |
| Db                    | 303             | AACTCTGTTTCTGCTCTTTGTGACTGCTATGATGCCAGACACCAACACTCACCATCTT      | 362       |             |
| Qy                    | 412             | CGTTCTGGACAAAGTATATGGAGATGCGCTGGAGGTTTGGGAGATTGTGA---GTTTTTCATC | 468       |             |
| Db                    | 363             | CGTATGGGATAAGCTTAAGAAAGAAGAACTTCCAGAGAGGAATGGTTAGACACCGTGTTTAC  | 422       |             |
| Qy                    | 469             | TCATGAACACAGAAATGGGACGATGAGTTTATTCGAAGATGGACCAAAATTCATGAAGC     | 528       |             |
| Db                    | 423             | GAGCGAGAAACCTGATTTTCTAGGAGTTTGGATTTTGGACCAAGAAATTCATGGCCAA      | 482       |             |
| Qy                    | 529             | ACGTCTTTTACAAAATTTGCCATAGAGGATTTAGCGCTGCCAAAAATTTGTTGCATAGGCA   | 588       |             |
| Db                    | 483             | GAACTTGTATCAGTTGTCTCCAGTCCAGATCTTGAATTCGCGAAAAATGTTGGTGAGGC     | 542       |             |
| Qy                    | 589             | AGGTCATTTTTACAGAGGATCTATCAAGAAAGAAAGTTTACCGAGGAGGATATGG         | 648       |             |
| Db                    | 543             | AAACCCCATTTGATTAAAGAAAGATATGGCAGAGAGAAAGCTTTCAGTGGAGGAGGATACGG  | 602       |             |
| Qy                    | 649             | TTCCGTTGCAACGAGTTTACGTTAATGAGTAGTGAAGACAAAGCCCATCCCTCGCATTTTCAT | 708       |             |

Db 603 ATCCGTTACAGTATATTATTATGTCGGAAGGATCTTGTGTACCCGAGATATACCA 662  
Qy 709 TCGTTGGATGATTGAATTTCAACGTCCTCGAAGTCTACGAGATCGATGGCGGATCA 768  
Db 663 GCATCGATGATCAGCAACTTTCCCCAAAAGAGTAATGGAGATCAAAAGACGAGATCA 722  
Qy 769 CATGGTGATGCTCTCAAAACCCCAAAACTCTTTGACTCTCTCTCTGCTATTGCCACCA 828  
Db 723 TATGCCAATGTTCTCCAGCCTCAACACTATGTGCTCTCTCTTGGAGATTGCAATAA 782  
Qy 829 TTATATGTAA 838  
Db 783 ATATGCCTAA 792

RESULT 10  
US-09-938-842A-217  
; Sequence 217, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRI1300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 217  
; LENGTH: 792  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-217

Query Match 24.0%; Score 218.8; DB 11; Length 792;  
Best Local Similarity 56.6%; Pred. No. 4.1e-58;  
Matches 447; Conservative 0; Mismatches 337; Indels 6; Gaps 2;  
Qy 55 GAGACATATGGAGAGAAACATCATCTCGTTGTTAGTTCAACAGCTTATCATGGAGCCTG 114  
Db 3 GAGTGAGGAGGAGAGAGCAACACGTCGTTCTAGTACATGGTGTGCTTGGCCATGGCGCTG 62  
Qy 115 GATCTGTTACAGCTCAAGCCCTCTTGAATCAGCGCGCCACCGGTTACTGCTGTGCA 174  
Db 63 GTGCTGTTAAGGTTAAGCCGAGCTCGAGGCTTCTGGCCACCGGTTAACCCGCGGTAGA 122  
Qy 175 ACTCGCGGCTCCGGATCGAC---CCACGACCAATCCAGGCGGTTGAAACCGTCGACGA 231  
Db 123 TCTAGTGTCTCGGTATAGACATGACACGAGTCAATCAGATATATCCAGATGCGAACA 182  
Qy 232 ATACTCCAAACCGTTGATCGAAACCCCTCAAACTCTTTCAGAGAACGAGAGTAATTC 291  
Db 183 ATACTCAGAGCAATTCATGACGCTAATGACCTCACTTACCAGATGATGAGAAGTTGTGCT 242  
Qy 292 GGTGATTCAGCTTCGGAGGATCAACATCGCTCTCGCGCGCGGACATATTTCCGCGGNA 351  
Db 243 TGTGTTCTAGCTTAGGAGGTTTGAATTTAGCTATGGCCATGGAATATTTTTCCGACCAA 302  
Qy 352 GATTAAGGTTCTTGCTTCTCAACGCTTCTTGGCCGACACACACCGCTGCTTCTCA 411  
Db 303 AATCTCTGTTCTGCTTCTGCTGCTGCTATGATGCGACACACCAACACTCAACATCTT 362  
Qy 412 CGTCTCGGACAAATATATGGAGATGCTCGAGGTTTGGGAGATTGTGA---GTTTTATC 468

Db 363 CGTATGGTAAGCTAAGAAAAGAAACTTTCAGAGAGGAATGGTTAGACACCGTGTTTAC 422  
Qy 469 TCATGAACAAGAAATGGACGATGAGTTTATTGAGATGGGACCAAAATTCATGAAGGC 528  
Db 423 GAGCGAGAAACCTGATTTTCTAGGAGTTTGGATTTTGGACCGAATTCATGGCCAA 482  
Qy 529 AGCTCTTTACCAAAATTTGCCATAGAGATTAACGAGCTGGCAAAATTTTCATAGGCA 588  
Db 483 GAACTTGATGATGTTCTCCAGTCCAAGATCTTGAATTTGGCGAAAATGTTGGTAGGCG 542  
Qy 589 AGGGTCATTTTTCAAGAGGATCTATCAAGAAAGAAAGTTTAGGAGGAAAGGATATGG 648  
Db 543 AAACCCATTTGATTAAAGAAAGATATGGCAGAGAAAGAGCTTCAGTAGAGAAAGGATACG 602  
Qy 649 TTCGGTGCACGAGTTTACGTAATCAGTAGTAGTGAACACAAAGCCATCCCTGCGATTTCA 708  
Db 603 ATCCGTTACACGATATATTATTGATGCGGAAGGATCTTGTGTACCCGGAAGATTACCA 662  
Qy 709 TCGTTGGATGATTGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGGATCA 768  
Db 663 GCGATCGATGATCAGCAACTTTCCCCCAAGAGAGTAATGGAGATCAAAAGACGAGATCA 722  
Qy 769 CATGTGATGCTCTCCAAACCCCAAAACTCTTTGACTCTCTCTGCTATTTGCCACCGA 828  
Db 723 TATGCCAATGTTCTCCAAAGCCTCAACAACTATGTGCTCTCTCTTGGAGATTGCAATAA 782  
Qy 829 TTATATGTAA 838  
Db 783 ATATGCCTAA 792

RESULT 11  
US-09-938-842A-1551  
; Sequence 1551, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRI1300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1551  
; LENGTH: 792  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1551

Query Match 23.5%; Score 214.2; DB 9; Length 792;  
Best Local Similarity 57.3%; Pred. No. 1.2e-56;  
Matches 450; Conservative 0; Mismatches 323; Indels 12; Gaps 3;  
Qy 55 GAGACATATGGAGAGAAACATCATCTCGTTGTTAGTTTCAACAGCTTATCATGGAGCCTG 114  
Db 3 GAGTGAGAGAGAGAAAGCAACATCTCGTGTAGTACATGGTGTGCGCGGCGCATG 62  
Qy 115 GATCTGTTACAGCTCAAGCCCTCTTGAATCAGCCGCGCCACCGCTTACTGCTGTGCA 174  
Db 63 GTGCTGTTAAGGTTAAGCCTCTTCTCGAGGCTTTGGGCCATCGTTAAACCGCCTTAGA 122  
Qy 175 ACTCGCGCCCTCCGGGATCGACCCA---CGACCAATCCAGGCGGTTGAAACCGTCGACGA 231  
Db 123 CCTAGCTGCTTCGGGTATAGACACAAACCGAGGTCAATCACTGACATTTCTTACATGTGAACA 182

```
QY 232 ATACTCCAAACCGTTGATCGAAACCCCTCAATCTCTTCAGAGAACGAGGTAATTTCT 291
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 ATATTTCTGAGCCATTGATGACGATTAATGATCTTCAATGCGGAATGATGAGAAGTTGTACT 242
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 292 GGTGGATTACGCTTCGAGGAGCATCAACATCGCTCTCGCCGCGGACATATTTCCGGCGAA 351
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 CGTTGGTCTAGCTTTGGAGGTTTGAGTTAGCTTAGCCATGGATAGTTTCCCGATAA 302
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 352 GATTAAGGTTCTGTGTTCTCAACGCTTCTTGGCCGACACAAACCCAGTGCCTTCTCA 411
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 AATCTCTGTCTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 412 CGTTCTGGCAAGTAT-----ATGAGATGCTCGAGGTTTGGAGATTGTGAGTTTTC 465
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 CGTCGAGGAAAGTTTGCAGAGCATGACACAGAGGATGGATGGGCTCTGAGCTCGA 422
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 466 ATCTCATGAAACAGAAATGGGACGATGATGATTTATTTGAAGATGGGACCAAAATTCATGAA 525
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 GACATATGGTTCAGATAATTCGGGCTTGTCTGTGTT---CTTCAGACCGACTTCATGAA 479
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 526 GGCAGCTCTTTACCAAAATGTCCTCATAGAGGATTCAGAGCTGGCAAAATGTTGCATAG 585
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 480 GCAACGCTCTACCAAACTTCTCTCTGAGGATCTTGAGCTTGGATTGCTTCTAAAGAG 539
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 586 GCAAGGTCATTTTTCACAGAGATCTATCAAGAAGAAAGTTTACGAGAGGATGAGGATA 645
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 540 GCCTAGTTCATTTTATTAATGAAATATCGAAGATGGAGAACTTTTCTGAGAAAGGTA 599
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 646 TGGTTCGGTCAACGAGTTTACGTAATGATGATGAGTGAAGCAAGCCATCCCTGCGATTT 705
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 600 TGAATCTGTCTCGAGCTTACATTTGTGTGTCGAAAGAGCAACATATCTCGGAAGACCA 659
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 706 CATTCGTTGGATGATGATTAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGAGA 765
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 660 TCAACGATGGATGATCCTAATTTATCCGCGCAATTTAGTGAATGAGATGGAAGAGACTGA 719
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 766 TCACATGGTGTGCTCTCAAAACCCCAAAACCTTTGACTCTCTCTCTCTCTCTCTCTCT 825
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 720 TCATATGCCAATGTTTTCGAAACCTCAACTACTAAGTGACCATCTATTGGCAATCGCTGA 779
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 826 CGATT 830
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 780 CAATT 784
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 12

```
US-09-938-842A-1551
; Sequence 1551, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/364,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1551
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1551
```

```
Query Match 23.5%; Score 214.2; DB 11; Length 792;
Best Local Similarity 57.3%; Pred. No. 1.2e-56;
Matches 450; Conservative 0; Mismatches 323; Indels 12; Gaps 3;
QY 55 GAGACATATGAGAGGAAACATCACTTCGTGTTAGTTTCAACAACGCTTATCATGAGGCTG 114
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3 GAGTGAGAGAGGAGGAAACAACTTCGTGCTAGTACATGGTGGTCCGACGCGCATG 62
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 115 GATCTGGTACAAAGCTCAAGCCCTTCCTTGAATCAGCCGCGCCACCGCTTACTGTGTGCA 174
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 63 GTGCTGGTACAAAGTTTAAAGCCTCTTCTCGAGGCTTTTGGGCGCATGCTGTAAACCGCTTAGA 122
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 175 ACTCGCGCTCTCGGATCGAACCA---CGACCAATCCAGCCGCTTGAACCGCTGACGA 231
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 CCTAGCTGCTTCCGGTATAGACACAAACAGGTCAATCACTGACATTTCTCATGTGAACA 182
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 232 ATACTCCAAACCGTTGATCGAAACCCCTCAATCTCTTCCAGAGAACGAGGTAATTTCT 291
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 ATATTTCTGAGCCATTGATGACGCTAATGACTTTCATTCGCGAATGATGAGAAGTTGTACT 242
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 292 GGTGGATTACGCTTCGAGGAGCATCAACATCGCTCTCGCCGCGGACATATTTCCGGCGAA 351
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 CGTTGGTCTAGCTTTGGAGGTTTGGAGTTTAGCTTAGCCATGGATAAGTTTCCCGATAA 302
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 352 GATTAAGGTTCTGTGTTCTCAACGCTTCTTCCCGACACAAACCCAGTGCCTTCTCA 411
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 AATCTCTGTCTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 412 CGTTCTGGCAAGTAT-----ATGAGATGCTCGAGGTTTGGAGATTGTGAGTTTTC 465
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 CGTCGAGGAAAGTTTGCAGACGATGACACAGAGGATGGATGGGCTCTGAGCTCGA 422
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 466 ATCTCATGAAACAGAAATGGGACGATGATGATTTATTTGAAGATGGGACCAAAATTCATGAA 525
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 GACATATGGTTCAGATAATTTCAACGCTTGTCTGTGTT---CTTCAGACCGACTTCATGAA 479
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 526 GGCAGCTCTTTACCAAAATGTCCTCATAGAGGATTCAGAGCTGGCAAAATGTTGCATAG 585
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 480 GCACGCTCTACCAACTTCTCTCTGAGGATCTTGAGCTTGGATTGCTTCTAAAGAG 539
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 586 GCAAGGTCATTTTTCACAGAGATCTATCAAGAAGAAAGTTTACGAGAGGATGAGGATA 645
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 540 GCCTAGTTCATTTTATTAATGAAATATCGAAGATGGAGAACTTTTCTGAGAAAGGTA 599
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 646 TGGTTCGGTCAACGAGTTTACGTAATGATGATGAGTGAAGCAAGCCATCCCTGCGATTT 705
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 600 TGAATCTGTCTCGAGCTTACATTTGTGTGTCGAAAGAGCAACATATCTCGGAAGACCA 659
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 706 CATTCGTTGGATGATGATTAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGCGAGA 765
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 660 TCAACGATGGATGATCCTAATTTATCCGCGCAATTTAGTGAATGAGATGGAAGAGACTGA 719
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 766 TCACATGGTGTGCTCTCAAAACCCCAAAACCTTTGACTCTCTCTCTCTCTCTCTCTCT 825
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 720 TCATATGCCAATGTTTTCGAAACCTCAACTACTAAGTGACCATCTATTGGCAATCGCTGA 779
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 826 CGATT 830
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 780 CAATT 784
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 13

```
US-10-780-002-33
; Sequence 33, Application US/10780002
; Publication No. US20050034196A1
; GENERAL INFORMATION:
; APPLICANT: Klessig, Daniel F.
; APPLICANT: Kumar, Dhirenda
; TITLE OF INVENTION: NOVEL SALICYLIC ACID-BINDING PROTEIN
; FILE REFERENCE: ENCODING NUCLEIC ACIDS, SAGEF2, AND METHODS OF USE THEREOF
; CURRENT APPLICATION NUMBER: US/10/780,002
; CURRENT FILING DATE: 2004-02-17
```

```
/ PRIOR APPLICATION NUMBER: PCT/US02/26312
/ PRIOR FILING DATE: 2002-08-16
/ PRIOR APPLICATION NUMBER: 60/312,863
/ PRIOR FILING DATE: 2001-08-16
/ NUMBER OF SEQ ID NOS: 58
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 33
/ LENGTH: 870
/ TYPE: DNA
/ ORGANISM: Arabidopsis Thaliana
US-10-780-002-33

Query Match      23.5%; Score 214.2; DB 21; Length 870;
Best Local Similarity 57.3%; Pred. No. 1.2e-56;
Matches 450; Conservative 0; Mismatches 323; Indels 12; Gaps 3;

QY 55 GAGACATATGGAGAGAAACATCATTGTTAGTTTCAACAGCTTATCATGGAGCCTG 114
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 23 GAGTGAAGGAAGGAAGCAACACTTCGTGCTAGTACATGGTGCCTGCACGGCGCATG 82
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 115 GATCTGTGTAACAAGCTCAAGCCCTCTTGAATCAGCGCGCCACCGGTTACTGCTGCGA 174
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 83 GTGCTGGTACAGGTTAAGCCTCTTCTCGAGGCTTTGGGCCATCGTGTAAACCGCCTTAGA 142
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 175 ACTCGCGCGCTCCGGATCGAACCA---CGACCAATCCAGGCGGTTTGAAACCGTCGACGA 231
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 143 CCTAGCTGCTTCGGGTATAGACAAACCAAGGTCAATCACTGACATTTCTACATGTGAACA 202
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 232 ATACTCAAACCGTTGATCGAAACCTCAAACTCTCTTCAGAGAACGAAGAGTAATTC 291
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 203 ATATTCTGAGCCATTGATGAGCTAATGACTTCAITTCGCCGAATGATGAGAAGGTTGTACT 262
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 292 GGTGTGATTGAGCTTCGGAGGATCAACATCGCTCTCGCGCGCGACATATTTCCGCGGAA 351
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 263 CGTTGTGATGATGTTGGAGGTTTGAGTTTAGCCTTAGCATGATGATGATTTCCCGATAA 322
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 352 GATTAAGGTTCTTGTGTTCTCAACGCTTCTTCCCGGACACCAACCCAGCTGCTTCTCA 411
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 323 AATCTGCTCTGCTCTGCTGACTGATTCATTCATGCGCGGACACCAACACATCAACCATGTT 382
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 412 CGTTCTGGACAAGTAT-----ATGGAGATGCTCGAGGTTTGGGAGATGTGAGTTTTC 465
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 383 CGTCGAGGAAAGTTTTCGACGAGCATGACACCAAGGAATGGATGGGCTCTGAGCTCGA 442
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 466 ATCTCATGAACAAGAAATGGGACGATGAGTTTATTTGAAGATGGGACCAAAATTCATGAA 525
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 443 GACATATGGTTAGATAATTCGGGCTTGTCTGTGTT---CTTCAGCACCAGCTTCATGAA 499
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 526 GGCAGCTCTTTACCAAAATTTGCCATAGAGGATTACGAGCTGGCAAAAATGTTGCATAG 585
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 500 GCACCGCTCTACCAACTTCTCTCTGGAGGATCTTGAGCTTGGATTGCTTCTAAAGAG 559
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 586 GCAAGGGTCATTTTTCACAGAGATCTATCAAGAAGAAAGTTTATAGCGAGAGAGATA 645
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 560 GCCTAGTTTCAATGTTTATTAATGAATATCGAAGATGGAGAACTTTCTGAGAAAGGTA 619
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 646 TGGTTCGCTGCAAGGTTTACGTAATGATGATGATGATGATGATGATGATGATGATGAT 705
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 620 TGGATCTGTCTCGAGCTTACATGTTGTGCAAGAGAGCAACATTAATTCGGAAGACCA 679
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 706 CATTCCTGTGGATGATTGATAATTTCAACGCTCGAAAGTCTACGAGATCGATGGCGGAGA 765
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 680 TCAACGATGATGATCAATTAATTCGGGGAATTTAGTATGATGATGATGATGATGATGAT 739
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 766 TCACATGGTGTGCTCTCOAAACCCCAAAATCTCTTTGATCTCTCTCTCTCTCTCTCTCT 825
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 740 TCATATGCCAAATGTTTTGCAAAACCTCAACTACTAAGTACGACCATCTATTGCGCAATCGCTGA 799
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 826 CGATT 830
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 800 CAATT 804
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 14
US-10-738-922-1
/ Sequence 1, Application US/10738922
/ Publication No. US20050032159A1
/ GENERAL INFORMATION:
/ APPLICANT: Ichige, Eita
/ APPLICANT: Semba, Hisashi
/ APPLICANT: Imanaka, Tadayuki
/ TITLE OF INVENTION: METHOD FOR LOW-TEMPERATURE CULTURE OF
/ TITLE OF INVENTION: MICROORGANISM
/ FILE REFERENCE: 08917-091001
/ CURRENT APPLICATION NUMBER: US/10/738,922
/ CURRENT FILING DATE: 2003-12-16
/ PRIOR APPLICATION NUMBER: JP 2002-365676
/ PRIOR FILING DATE: 2002-12-17
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 777
/ TYPE: DNA
/ ORGANISM: Manihot esculenta Crantz
US-10-738-922-1

Query Match      22.5%; Score 205.4; DB 21; Length 777;
Best Local Similarity 55.3%; Pred. No. 7.2e-54;
Matches 420; Conservative 0; Mismatches 336; Indels 3; Gaps 1;

QY 77 CACTTCGTGTAGTTTCACAAACGCTTATCATGAGCGCTTGGATCTGGTACAGCTCAAGCCC 136
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 13 CATTTTGTGTGATTTCATACCATTTTGCATGCTGATGGATTTGGCATTAAGCTCAAAACA 72
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 137 CTCCTTGAATCAGCGCGGCACCGCGTACTGCTGTCGAACCTCGCGCCTCCGGGATCGAC 196
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 73 GCCTTTGAGAGAGCTGGCCCAAAAGTCACTGCACTGCACTGCACTGCACTGCACTGCA 132
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 197 CCACGACCAATCCAGGCGGTTGAAACCGTCGACGAATCTCCAAACCGTTGATCGAAACC 256
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 133 CCAGGCAAAATGAGCAGATTAATTCATTTGATGAATACTCTGAACCCCTTATTGACTTTC 192
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 257 CTCAAATCTCTCCAGAGAACGAAGAGTAACTCTGTTGGATTGAGCTTCGGAGGCATC 316
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 193 TTGGAGAAATCCCTCAAGGGGAAAGGTCATCTTGTGGTGAGAGCTGTGAGGGCTG 252
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 317 AACATCGCTCTCGCGCGGACATATTTCCGGGGAAGATTAAGGTTCTTGTGTTCTCTCAAC 376
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 253 AATATTGCTATTGCTGCTGATAGATACGTTGACAAATTTGCAGCTGTTGTTTCCAAAT 312
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 377 GCCTTCTTCCGCGACACAAACCCAGCTTCTCAAGTTCTGGAACAAGTATATGAGATG 436
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 313 TCCTTTATGGCAGACACCGTTTCATAGCCCATCTTACACTGTGGAAGAGCTTTTGGAGTCG 372
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 437 CCGTGGAGTTTGGGAGATTGTGAGTTTTCATCTCATGAAACAAGAAATGG---GACGATG 493
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 373 TTTCTGATCGAGAGACACAGAGTATTTTACGTTTCTCTAATATCATCTGGAGAGACAAT 432
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 494 AGTTTATTTGAAGATCGGACCAAAATTCATGAAGGACGCTTTTACCAAAATTTGCCATA 553
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 433 ACAACATGAGCTGGCTTCGTACTTCTGAGGGAATTTATTTTACCAATTTGCATGAT 492
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 554 GAGGATTCAGAGCTGGCAAAAATGTTGATAGGCAAGGATCAATTTTTCACAGAGGATCTA 613
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 493 GGGGAATATGAATCGCAAAAATGTAATAGGAAGGGATCACTGTTTCAAAATGTTTGG 552
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 614 TCAAGAAGAAAGATTTAGCGAGGAAGATATGTTTGGTCCAGGATTTACGTAATG 673
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 553 GCTCAGAGACGGAAGTTCCACGAAAGGTTACCGGATCAATTAAGAAAGTTTATATTGG 612
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 674 AGTAGTGAAGACAAAGCCATCCCTGCGATTTTCACTTCTGTTGATGATTAATTTCAAC 733
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 613 ACCGATCAAGACAAATATTTTACCAGACTTTTCAACGCTGGCAATTTGCAAACTACAAA 672
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 734 GTCTCGAAGTCTACGAGATCGATCGCGGAGATCAATGGTGATGCTCTTCCAAAACCCCAA 793
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 673 CCAGACAGGTTTATCAGGTTCAAGTGGAGATCATAGCTCCAGCTTACAAAACTGAG 732

Qy 794 AAACCTCTTTGACTCTCTCTCTGCTATTTGCCACCGATTAT 832

Db 733 GAGGTAGCTCATATTCTCCAGAGGTGGCTGATGCATAT 771

RESULT 15

US-10-738-927-1

; Sequence 1, Application US/10738927

; Publication No. US20050032191A1

; GENERAL INFORMATION:

; APPLICANT: Semba, Hisashi

; APPLICANT: Ichige, Eita

; APPLICANT: Mukouyama, Masaharu

; TITLE OF INVENTION: METHOD FOR PRODUCTION OF

; TITLE OF INVENTION: S-HYDROXYNITRILE LYASE BY USE OF ESCHERICHIA COLI

; FILE REFERENCE: 08917-092001

; CURRENT APPLICATION NUMBER: US/10/738,927

; CURRENT FILING DATE: 2003-12-16

; PRIOR APPLICATION NUMBER: JP 2002-365675

; PRIOR FILING DATE: 2002-12-17

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 777

; TYPE: DNA

; ORGANISM: Manihot esculenta Crantz

US-10-738-927-1

Query Match 22.5%; Score 205.4; DB 21; Length 777;

Best Local Similarity 55.3%; Pred. No. 7.2e-54;

Matches 420; Conservative 0; Mismatches 336; Indels 3; Gaps 1;

Qy 77 CACTTCGTGTAGTTCCACAACGCTTATCATGTGAGCTGGATCTGGTACAAAGCTCAAGCC 136

Db 13 CATTTTGTCTGATTCATACCATTTGCCATGGTGTGATGGATTTGGCATRAGCTCAACCA 72

Qy 137 CTCCTTGAATCAGCGCGCCACCGGTTACTGCTGTGAACTCGCCGCTCCGGGATCGAC 196

Db 73 GCCCTTGAGAGAGCTGGCCACAAAGTCACTGCACATGGCAGCGGCGCATTTGAC 132

Qy 197 CCAGCACCAATCCAGCGCGTTGAAACCGTCGAGGATACCTCCAAACCGTTGATCGAAACC 256

Db 133 CCAAGGCAAAATTTGACAGATTAATTCATTTGATGAATACTCTGAACCCCTTATTGACTTTC 192

Qy 257 CTCAAATCTCTCCAGAGAACGAAGAGTAAATCTGTTGGATTTCAGCTTCGAGGCGATC 316

Db 193 TTGGAGAAACTCCTCAAGGGGAAAAGGTCACTATTGTTGGTGAGAGCTGTGCAGGGCTG 252

Qy 317 AACATGCTCTCCCGCCGACATATTTCCGGCGAAGATTAAAGTTCTTTGTGTTCTCTCAAC 376

Db 253 AATATTGCTATTGCTGCTGATAGATACGTTGACAAAATTCAGCTGGTGGTGTTCACCAAT 312

Qy 377 GCCTTCTCCCGACACACACCCACGTCCTTCTCAGTTCTCGACAAGTATATGGAGATG 436

Db 313 TCCCTTATGCCAGACACCGTTTATAGCCCATCTTACACTGTGGAAGAGCTTTTGGAGTCG 372

Qy 437 CTGGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAAACAAGAAATGG---GACGATG 493

Db 373 TTTCTGACTGGAGAGACACAGATATTTTAGCTTCACTAATATCACTGGAGAGACAAAT 432

Qy 494 AGTTTATCGAATGGGACCAAAATTCATGAAGGCGCTCTTTTACAAAATTTGTCCCATATA 553

Db 433 ACAACAATGAAGCTGGGCTTCGTACTCTGAGGGGAAAATTTATTTACCAAAATGCCACTGAT 492

Qy 554 GAGGATTACAGCTGGCAAAAATTTGTCATAGCCAGGGTCAATTTTTCACAGAGGATCTA 613

Db 493 GGGGATATGAACTGGCAAAAATTTGGAATGAGGAAGGGATCACTGTTCAAAATGTTTGG 552

Qy 614 TCAAGAGAAAGAAAGTTTAGCGAGGAAGGATATGTTTCGGTCAACAGGTTTACGTAATG 673

Db 553 GCTCAGAGACCGNAGTTCCACGNAAGGTTACGGATCAATTAAGAAAGTTTATATTGG 612

Qy 674 AGTAGTGAAGACAAAGCCATCCCTGCGATTTCATTCGTTGGATGATGATATTTCAAC 733

Db 613 ACCGATCAAGACAAAATATTTTACCAGACTTTTCAACGCTGGCAAATTTGCAAACCTACAAA 672

Qy 734 GTCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTGGTGTCTCTCCAAACCCCAA 793

Db 673 CCAGACAAAGGTTTATCAGGTTCAAGGTGGAGATCATAGCTCCAGCTTACAAAACCTGAG 732

Qy 794 AAACCTCTTTGACTCTCTCTCTCTGCTATTTGCCACCGATTAT 832

Db 733 GAGGTAGCTCATATTCTCCAAGAGGTGGCTGATGCATAT 771

Search completed: September 12, 2005, 12:48:37

Job time : 2023.05 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 00:50:21 ; Search time 2907.77 Seconds  
(without alignments)  
11938.557 Million cell updates/sec

Title: US-10-780-002-36

Perfect score: 912

Sequence: 1 atcagctattctcaacaa.....ttttccgggcaactttcatc 912

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hc:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gss1:\*

9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 827   | 90.7        | 865    | CNS0926U | Arabidops          |
| 2          | 662   | 72.6        | 667    | AV822206 | AV822206           |
| 3          | 624   | 68.4        | 625    | AV785311 | AV785311           |
| 4          | 603.4 | 66.2        | 641    | CB258477 | CB258477 05-801273 |
| 5          | 557   | 61.1        | 572    | AV822640 | AV822640           |
| 6          | 533.8 | 58.5        | 537    | AV783452 | AV783452           |
| 7          | 434.8 | 47.7        | 441    | AV824147 | AV824147           |
| 8          | 421   | 46.2        | 421    | BP613009 | BP613009           |
| 9          | 418.6 | 45.9        | 480    | AA04853  | 28251 Lam          |
| 10         | 397.4 | 43.6        | 425    | BP642196 | BP642196           |
| 11         | 391.2 | 42.9        | 426    | AV790585 | AV790585           |
| 12         | 374.2 | 41.0        | 390    | AV820476 | AV820476           |
| 13         | 366   | 40.1        | 366    | AV554161 | AV554161           |
| 14         | 359.8 | 39.5        | 405    | BP642613 | BP642613           |
| 15         | 357.8 | 39.2        | 361    | AI933246 | AI933246           |
| 16         | 357.6 | 39.2        | 382    | BP646852 | BP646852           |
| 17         | 298.4 | 32.7        | 335    | AV545429 | AV545429           |
| 18         | 282.8 | 31.0        | 351    | BP604502 | BP604502           |
| 19         | 259.2 | 28.4        | 797    | CF837566 | CF837566           |
| 20         | 247   | 27.1        | 308    | AV559909 | AV559909           |
| 21         | 246.8 | 27.1        | 797    | CB290370 | CB290370           |
| 22         | 246.8 | 27.1        | 804    | CB293837 | CB293837           |
| 23         | 246.8 | 27.1        | 821    | CB290604 | CB290604           |
| 24         | 246.8 | 27.1        | 865    | CN190989 | CN190989           |

|    |       |      |     |   |          |           |
|----|-------|------|-----|---|----------|-----------|
| 25 | 246   | 27.0 | 781 | 7 | CN190896 | CN190896  |
| 26 | 244.8 | 26.8 | 850 | 6 | CB292674 | UCRCS06_0 |
| 27 | 243.6 | 26.7 | 834 | 7 | CN191164 | UCRCS06_0 |
| 28 | 243.2 | 26.7 | 811 | 6 | CB292675 | UCRCS01_0 |
| 29 | 242   | 26.5 | 828 | 7 | CF838704 | UCRCS03_0 |
| 30 | 239.8 | 26.3 | 751 | 7 | CN191146 | UCRCS06_0 |
| 31 | 239   | 26.2 | 835 | 7 | CF830288 | UCRCS01_0 |
| 32 | 237   | 26.0 | 747 | 7 | CN189384 | UCRCS06_0 |
| 33 | 234.2 | 25.7 | 556 | 8 | BZ507980 | BONAQ86FR |
| 34 | 232.4 | 25.5 | 793 | 7 | CF838361 | UCRCS03_0 |
| 35 | 230.6 | 25.3 | 740 | 7 | CN191507 | UCRCS06_0 |
| 36 | 228.8 | 25.1 | 726 | 7 | CF831872 | UCRCS01_0 |
| 37 | 228.2 | 25.0 | 726 | 7 | CF831499 | UCRCS01_0 |
| 38 | 225.8 | 24.8 | 957 | 7 | CV267710 | WS02032_B |
| 39 | 219.2 | 24.0 | 911 | 3 | CNS0A981 | Arabidops |
| 40 | 217   | 23.8 | 217 | 1 | AV782897 | Arabidops |
| 41 | 214.4 | 23.5 | 907 | 7 | CK266902 | EST712980 |
| 42 | 214.2 | 23.5 | 886 | 7 | CK260266 | EST706344 |
| 43 | 214.2 | 23.5 | 914 | 7 | CK270870 | EST716948 |
| 44 | 214.2 | 23.5 | 966 | 7 | CK260265 | EST706343 |
| 45 | 213   | 23.4 | 719 | 7 | CN190831 | UCRCS06_0 |

ALIGNMENTS

RESULT 1  
CNS0926U  
LOCUS  
DEFINITION  
CNS0926U  
865 bp mRNA linear HTC 06-FEB-2004  
Arabidopsis thaliana Full-length cDNA complete sequence from clone GSLTSL1782C11 of Silique of strain col-0 of Arabidopsis thaliana (thale cress).  
ACCESSION  
BX833804  
VERSION  
BX833804.1  
KEYWORDS  
HTC; GSLT cDNA.  
SOURCE  
Arabidopsis thaliana (thale cress)  
ORGANISM  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
REFERENCE  
Castelli V., Aury J.M., Jallou O., Wincker P., Clepet C., Menard M., Cruaud C., Quetier F., Scarpelli C., Schachter V., Temple G., Caboche M., Weissenbach J., and Salanoubat M.  
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation  
JOURNAL  
Unpublished  
REFERENCE  
2 (bases 1 to 865)  
AUTHORS  
Direct Submission  
JOURNAL  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
COMMENT  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C., URGV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full-length  
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.  
location/Qualifiers  
1..865  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"

FEATURES  
source







Db 422 CAAGTATATGGAGATCCCTGGAGGTTTGGAGATTCTGAGTTTTCATCTCATGAACAAG 481  
 Qy 481 AATATGGAGATGAGTTTATTTGAAGATGGACCAAAATTCATGAAGGACGCTTTTACCA 540  
 Db 482 AATATGGAGATGAGTTTATTTGAAGATGGACCAAAATTCATGAAGGACGCTTTTACCA 541  
 Qy 541 AATATGTCCTCATAGAGATTACGAGCTGGCAAAATGTTGTCATAGGCAAGGTCATTTT 600  
 Db 542 AATATGTCCTCATAGAGATTACGAGCTGGCAAAATGTTGTCATAGGCAAGGTCATTTT 601  
 Qy 601 CACAGAGATCTATCAAGAAAGAAAGTTTACGAGGAGGATATGTTTCCGTCGAACG 660  
 Db 602 CACAGAGATCTATCAAGAAAGAAAGTTTACGAGGAGGATATGTTTCCGTCGAACG 661  
 Qy 661 AGTT 664  
 Db 662 AGTT 665

## RESULT 3

AV785311/c 625 bp mRNA linear EST 28-MAR-2002

LOCUS AV785311 RAFL6 Arabidopsis thaliana cDNA clone RAF106-14-F12 3', mRNA sequence.

DEFINITION AV785311 GI:19804101

ACCESSION AV785311.1

VERSION Arabidopsis thaliana (thale cress)

KEYWORDS Arabidopsis thaliana

SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana

REFERENCE Arabidopsis thaliana

AUTHORS Arabidopsis thaliana

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 625)

Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,

Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,

Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.

and Shinozaki, K.

Large scale analysis of Arabidopsis full-length cDNA (2002b)

Unpublished (2002)

Contact: Motoaki Seki

Plant Functional Genomics Research Group

RIKEN Genomic Sciences Center

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: 81-298-36-4359

Fax: 81-298-36-9060

Email: msekic@rtc.riken.go.jp

An Arabidopsis full-length cDNA library was constructed essentially

as reported previously (Seki et al., 1998). This clone is in a

modified pBluescript vector as a SstI/XhoI insert. Please visit our

web site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for

further details.

Location/Qualifiers

1..625

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/clone="RAFL06-14-F12"

/dev\_stage="plants at various developmental stages from

germination to mature seeds"

/lab\_host="DH10B"

/clone\_lib="RAFL6"

/note="Site 1: SstI; Site 2: XhoI; subjected to

dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24

hr) treatments"

ORIGIN

Query Match 68.4%; Score 624; DB 1; Length 625;

Best Local Similarity 100.0%; Pred. No. 1.5e-179;

Matches 624; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 279 AAGAGTATCTCGTTGGATTGAGTTCGAGGAGCATCAACATCGTCTCGCGCCGACG 338

Db 625 AAGAGTATCTCGTTGGATTGAGTTCGAGGAGCATCAACATCGTCTCGCGCCGACG 566  
 Qy 339 TATTTCCGCGCAAGATTAAAGTTCTTGTGTTCTCAAGCCCTTCTTGGCCGACACACCC 398  
 Db 565 TATTTCCGCGCAAGATTAAAGTTCTTGTGTTCTCAAGCCCTTCTTGGCCGACACACCC 506  
 Qy 339 ACGTGCTTCTCAAGTATATGGAGATCGCTTGGAGATCGCTTGGAGATCGCTTGGAGATCG 458  
 Db 505 ACGTGCTTCTCAAGTATATGGAGATCGCTTGGAGATCGCTTGGAGATCGCTTGGAGATCG 446  
 Qy 459 AGTTTTCATCATGAACAAAGAAATGGAGATCGATGATTTATTAAGATGGAGACCAAAAT 518  
 Db 445 AGTTTTCATCATGAACAAAGAAATGGAGATCGATGATTTATTAAGATGGAGACCAAAAT 386  
 Qy 519 TCATGAAGGACGCTCTTTTACCAAAATCTCCCATAGAGATTTACGAGCTGGCAAAATGCT 578  
 Db 385 TCATGAAGGACGCTCTTTTACCAAAATGCTCCCATAGAGATTTACGAGCTGGCAAAATGCT 326  
 Qy 579 TGCATAGGCAAGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTACGAGG 638  
 Db 325 TGCATAGGCAAGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTACGAGG 266  
 Qy 639 AAGGATATGTTTCGTCGACAGGTTTACGTAATGAGTAGTAGAGACAAAGCCATCCCT 698  
 Db 265 AAGGATATGTTTCGTCGACAGGTTTACGTAATGAGTAGTAGAGACAAAGCCATCCCT 206  
 Qy 699 GCGATTTTCATTCGTTGGAGATTTCAAGCTCTCGAAAGCTCTCGAAAGCTCTCTCTGCTA 758  
 Db 205 GCGATTTTCATTCGTTGGAGATTTCAAGCTCTCGAAAGCTCTCTCTGCTA 146  
 Qy 759 GCGAGATCATCATGATGATGCTCTCCAAACCCCAAAACTCTTTGACTCTCTCTGCTA 818  
 Db 145 GCGAGATCATCATGATGATGCTCTCCAAACCCCAAAACTCTTTGACTCTCTCTGCTA 86  
 Qy 819 TTGCCACCGATTATGTAATTAATCTTAAGTCCGTTTACTTTTCTCATCGTTACTAA 878  
 Db 85 TTGCCACCGATTATGTAATTAATCTTAAGTCCGTTTACTTTTCTCATCGTTACTAA 26  
 Qy 879 TAAACAAACCCCTTTTTCGCGGC 902  
 Db 25 TAAACAAACCCCTTTTTCGCGGC 2

## RESULT 4

CB258477

LOCUS

DEFINITION

CB258477

ACCESSION

CB258477

VERSION

CB258477.1

GI:32883250

EST

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 641)

Schmid, K.J., Soerensen, T.R., Stracke, R., Torjek, O., Altmann, T.,

Mitchell-Olds, T. and Weishaar, B.

Large-scale identification and analysis of genome-wide

single-nucleotide polymorphisms for mapping in Arabidopsis thaliana

Genome Res. 13 (6), 1250-1257 (2003)

22683290

12799357

COMMENT

Contact: Weishaar B

ADIS DNA core facility at MPIZ

Max-Planck-Institute for Plant Breeding Research

Carl-von-Linne Weg 10, 50829 Koeln, Germany

Fax: 00492215062851

Email: weishnaa@mpiz-koeln.mpg.de

Insert Length: 641 Std Error: 0.00

Plate: 2 row: J column: 01

Seq primer: T7R; CTAATACGACTCTACTATAGGA.

|           |   |  |
|-----------|---|--|
| FEATURES  | Location/Qualifiers   |  |
|           | 1..641  |  |
| source    | /organism="Arabidopsis thaliana"  |  |
|           | /mol_type="mRNA"  |  |
|           | /cultivar="Cape Verde Island (Cvi)"   |  |
|           | /db_xref="GABI:590641"  |  |
|           | /db_xref="taxon:3702"   |  |
|           | /clone="MP12p771J012Q"  |  |
|           | /issue_type="whole plant"   |  |
|           | /dev_stages="adult plant, mixed stresses"   |  |
|           | /lab_host="E. coli TOP10"   |  |
|           | /clone_lib="MP12-ADIS-014"  |  |
|           | /note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; cDNA library from Arabidopsis thaliana, accession Cvi; six weeks old total plants grown under long-day conditions in soil. whole adult plants were treated for 24 hours with different stresses, (1) at 4 Grad C in the dark, (2), at 37 Grad C in the dark, (3) lying in the lab after removing from soil, (4) in the greenhouse after wounding leaves with a foreceps, (5) in the lab watering with a 150 mM NaCl solution, (6) at 26 Grad C in the light/UV; equal quantities of stressed plant material were pooled; library was made at the Max-Planck-Institute for Plant Breeding Research, Cologne, Germany; cloning sites SalI-NotI, primer sites and orientation: |  |
|           | T7-SalI-CCACGGCTCCG-Sprime-cDNA-polyA-CC-NotI-SP6; Note: Sequencing granted in the context of the GABI Arabidopsis Verbund I: Genetic Diversity, 'Establishment of high-efficiency SNP-based mapping tools and development of methods for genome-wide mutation detection' PI: Bernd Weisshaar Sequence submission managed by RZPD/GABI-Primary database: <a href="http://gabi.rzpd.de">http://gabi.rzpd.de</a> . This clone is available from RZPD; contact RZPD ( <a href="mailto:clone@rzpd.de">clone@rzpd.de</a> ) for further information."   |  |
| ORIGIN    | Query Match 66.2%; Score 603.4; DB 6; Length 641;   |  |
|           | Best Local Similarity 99.8%; Pred. No. 3.2e-173;  |  |
|           | Matches 604; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  |  |
|           |   |  |
| QY        | 25  | AGAACCTTAACCAAGAAACGCTACGGAGACATATGGAGGAAACATCACTTCGT 84           |
| DB        | 37  | AGAACCTTAACCAAGAAACGCTACGGAGACATATGGAGGAAACATCACTTCGT 96           |
| QY        | 85  | GTTAGTTACAAACGCTTATCATGAGCGCTGGATCTGTGTAAGCTCAAGCCCTCTTGA 144      |
|           | 97  | GTTAGTTACAAACGCTTATCATGAGCGCTGGATCTGTGTAAGCTCAAGCCCTCTTGA 156      |
| QY        | 145   | ATCAGCGGCCACCGGTTACTGTGTGTAACCTCGCGGCTCCGGGATCGACCCAGACC 204       |
| DB        | 157   | ATCAGCGGCCACCGGTTACTGTGTGTAACCTCGCGGCTCCGGGATCGACCCAGACC 216       |
| QY        | 205   | AATCCAGGCGTTGAACCGTCGACGATACTCCAAACCGTTGATCGAAACCTCAAAATC 264      |
| DB        | 217   | AATCCAGGCGTTGAACCGTCGACGATACTCCAAACCGTTGATCGAAACCTCAAAATC 276      |
| QY        | 265   | TCTTCAGAGAACGAAGGTAATTCTGGTTGGATTTCAGCTTCGGAGGCATCAACATCGC 324     |
| DB        | 277   | TCTTCAGAGAACGAAGGTAATTCTGGTTGGATTTCAGCTTCGGAGGCATCAACATCGC 336     |
| QY        | 325   | TCTCGCGCCGACATATTTCGCGCGAAGATTAAAGTTCTTGTTGTTCTCAAGCCCTTCTT 384    |
| DB        | 337   | TCTCGCGCCGACATATTTCGCGCGAAGATTAAAGTTCTTGTTGTTCTCAAGCCCTTCTT 396    |
| QY        | 385   | GCCCGACACACCGCTCTTCTCAAGTTCTGGACAGTATATGGAGATGCTCTGGAGG 444        |
| DB        | 397   | GCCCGACACACCGCTCTTCTCAAGTTCTGGACAGTATATGGAGATGCTCTGGAGG 456        |
| QY        | 445   | TTTGGAGATTGTGAGTTTCTATCTCATGAACAAAGAAATGGGACGATGAGTTTATGAA 504     |
| DB        | 457   | TTTGGAGATTGTGAGTTTCTATCTCATGAACAAAGAAATGGGACGATGAGTTTATGAA 516     |
| QY        | 505   | GATGGGACCAAAATTGATGAAGGCGAGTCTTTACCAAAATTTGTCCTCATAGAGGATTACGA 564 |
| Db        |   |  |
|           | 517   | GATGGGACCAAAATTGATGAAGGCGAGTCTTTACCAAAATTTGTCCTCATAGAGGATTACGA 576 |
| QY        |   |  |
|           | 565   | GCTGCGCAAAATGTTGCATAGGCAAGGTCATTTTTTTCACAGAGGATCTATCAAAAGAAAGA 624 |
| DB        |   |  |
|           | 577   | GCTGCGCAAAATGTTGCATAGGCAAGGTCATTTTTTTCACAGAGGATCTATCAAAAGAAAGA 636 |
| QY        |   |  |
|           | 625   | AAAGT 629  |
| DB        |   |  |
|           | 637   | AAAGT 641  |
| RESULT 5  | AV822640 572 bp mRNA linear EST 01-APR-2002   |  |
|           | LOCUS AV822640 RAF15 Arabidopsis thaliana cDNA clone RAFL05-09-N03 5', mRNA sequence.   |  |
| ACCESSION | AV822640  |  |
|           | VERSION AV822640.1 GI:19864692  |  |
| KEYWORDS  | EST.  |  |
|           | SOURCE Arabidopsis thaliana (thale cress)   |  |
| ORGANISM  | Arabidopsis thaliana  |  |
|           | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.   |  |
| REFERENCE | 1 (bases 1 to 572)  |  |
|           | Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.   |  |
| AUTHORS   | Large scale analysis of Arabidopsis full-length cDNA (2002b)  |  |
|           | Unpublished (2002)  |  |
| JOURNAL   | Contact: Motoaki Seki   |  |
|           | Plant Functional Genomics Research Group  |  |
| COMMENT   | RIKEN Genomic Sciences Center   |  |
|           | 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan   |  |
|           | Tel: 81-298-36-4359   |  |
|           | Fax: 81-298-36-9060   |  |
|           | Email: <a href="mailto:maeki@rtc.riken.go.jp">maeki@rtc.riken.go.jp</a>   |  |
|           | An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a modified pBluescript vector as a SstI/XhoI insert. Please visit our web site ( <a href="http://www.gsc.riken.go.jp/e/plant/index_e.html">http://www.gsc.riken.go.jp/e/plant/index_e.html</a> ) for further details.  |  |
| FEATURES  | Location/Qualifiers   |  |
|           | 1..572  |  |
| source    | /organism="Arabidopsis thaliana"  |  |
|           | /mol_type="mRNA"  |  |
|           | /db_xref="taxon:3702"   |  |
|           | /clone="RAFL05-09-N03"  |  |
|           | /dev_stages="rosette plants"  |  |
|           | /lab_host="SOLR"  |  |
|           | /clone_lib="RAFL5"  |  |
|           | /note="Site 1: SstI; Site 2: XhoI; subjected to dehydration-treated (1,2,5,10,24 hr)"   |  |
| ORIGIN    | Query Match 61.1%; Score 557; DB 1; Length 572;   |  |
|           | Best Local Similarity 99.5%; Pred. No. 5.2e-159;  |  |
|           | Matches 568; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  |  |
|           |   |  |
| QY        | 1   | ATCAGCGTATTCTCAACAGCAAGAACCTTACCAAAAGAAACGCTCTACGGAGACA 60         |
| DB        | 2   | ATCAGCGTATTCTCAACAGCAAGAACCTTACCAAAAGAAACGCTCTACGGAGACA 61         |
| QY        |   |  |
|           | 61  | TATGAGAGGAAACATCACTTCGTTGTTAGTTCAACAGCTTATCATGAGGCTTGATCTG 120     |
| DB        |   |  |
|           | 62  | TATGAGAGGAAACATCACTTCGTTGTTAGTTCAACAGCTTATCATGAGGCTTGATCTG 121     |
| QY        |   |  |
|           | 121   | GTACAAGCTCAAGCCCTCTCTTGAATCAGCGGCCACCGCGTTACTGCTGCGAATCTGCG 180    |
| DB        |   |  |
|           | 122   | GTACAAGCTCAAGCCCTCTCTTGAATCAGCGGCCACCGCGTTACTGCTGCGAATCTGCG 181    |

QY 181 CGCCTCCGGGATCGACCCACGACCAATCCAGGCGCTTGAAACCGTCGACGAATACTCCAA 240  
 Db 182 CGCCTCCGGGATCGACCCACGACCAATCCAGGCGCTTGAAACCGTCGACGAATACTCCAA 241  
 QY 241 ACGTTTGATCGAAACCTCAATCTCTTCAGAGAACGAGGTAATTCTGG- TTGGAT 299  
 Db 242 ACGTTTGATCGAAACCTCAATCTCTTCAGAGAACGAGGTAATTCTGGTTGGAT 301  
 QY 300 TCAGCTTCGGAGGATCAACATCGCTCTCGCGCGACACATATTTCCGGCGAAGATTAAAGG 359  
 Db 302 TCAGCTTCGGAGGATCAACATCGCTCTCGCGCGACACATATTTCCGGCGAAGATTAAAGG 361  
 QY 360 TTCTTGTTCTCCAAACGCTTTCTGCGCGACACACCCACGTCCTTCACGTTCTGG 419  
 Db 362 TTCTTGTTCTCCAAACGCTTTCTGCGCGACACACCCACGTCCTTCACGTTCTGG 421  
 QY 420 ACAAGTATATGAGATGCTTCGAGGTTTGGAGATTGTGAGTTTCATCTCATGAACAA 479  
 Db 422 ACAAGTATATGAGATGCTTCGAGGTTTGGAGATTGTGAGTTTCATCTCATGAACAA 481  
 QY 480 GAAATGGGACGATGAGTTTATTGAAGATGGAGCCAAATTCATGAAGGCAAGCTTTTACC 539  
 Db 482 GAAATGGGACGATGAGTTTATTGAAGATGGAGCCAAATTCATGAAGGCAAGCTTTTACC 541  
 QY 540 AAAATTGTCCCATAGAGGATTACGAGCTGGC 570  
 Db 542 AAAATTGTCCCATAGAGGATTACGAGCTGNC 572

RESULT 6  
 AV783452/c  
 LOCUS AV783452 RAFL5 Arabidopsis thaliana CDNA clone RAFL05-09-N03 3',  
 DEFINITION AV783452 537 bp mRNA linear EST 28-MAR-2002  
 mRNA sequence.  
 ACCESSION AV783452  
 VERSION AV783452.1 GI:19802242  
 KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 537)  
 Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,  
 Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,  
 Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.  
 and Shinozaki, K.  
 Large scale analysis of Arabidopsis full-length cDNA (2002b)  
 Unpublished (2002)  
 Contact: Motoaki Seki  
 Plant Functional Genomics Research Group  
 RIKEN Genomic Sciences Center  
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
 Tel: 81-298-36-4359  
 Fax: 81-298-36-9060  
 Email: msek@rtc.riken.go.jp  
 An Arabidopsis full-length cDNA library was constructed essentially  
 as reported previously (Seki et al., 1998). This clone is in a  
 modified pBluescript vector as a SstI/XhoI insert. Please visit our  
 web site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for  
 further details.

TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)

JOURNAL  
 COMMENT  
 FEATURES  
 source  
 1..537  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="RAFL05-09-N03"  
 /dev\_stage="rosette plants"  
 /lab\_host="SOLa"  
 /clone\_lib="RAFL5"  
 /note="Site 1: SstI; Site 2: XhoI; subjected to  
 dehydration-treated(1,2,5,10,24 hr)"

ORIGIN

Query Match 58.5%; Score 533.8; DB 1; Length 537;  
 Best Local Similarity 99.6%; Pred. No. 6e-152;  
 Matches 535; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 376 GCCTTCCTTGCCCGACACACCCACGTCCTTCACGTTCTGGCAAGTATATGAGAT 435  
 Db 537 GCCTTCCTTGCCCGACACACCCACGTCCTTCACGTTCTGGCAAGTATATGAGAT 478  
 QY 436 GCCTGGAGGTTGGGAGATTGTGAGTTTTCATCTCATGAACAGAAATGGGACCATGAG 495  
 Db 477 GCCTGGAGGTTGGGAGATTGTGAGTTTTCATCTCATGAACAGAAATGGGACCATGAG 418  
 QY 496 TTATTGAAGATGGGACCAAAATTCATGAAGCAGCTCTTTACCAAAAATTTGCCATAGA 555  
 Db 417 TTATTGAAGATGGGACCAAAATTCATGAAGCAGCTCTTTACCAAAAATTTGCCATAGA 358  
 QY 556 GGATTACGAGCTGGCAAAAATTTGTCATAGCAAGGGTCATTTTTCACAGAGGATCTATC 615  
 Db 357 GGATTACGAGCTGGCAAAAATTTGTCATAGCAAGGGTCATTTTTCACAGAGGATCTATC 298  
 QY 616 AAAGAAAGAAAGTTTAGCGGAGGAGGATATGTTTCGGTGCACAGGTTTACGTAATGAG 675  
 Db 297 AAAGAAAGAAAGTTTAGCGGAGGAGGATATGTTTCGGTGCACAGGTTTACGTAATGAG 238  
 QY 676 TAGTGAACACAAAGCCATCCCTCGGATTTTCATTCGTTGGATGATTGATAATTTCAACGT 735  
 Db 237 TAGTGAACACAAAGCCATCCCTCGGATTTTCATTCGTTGGATGATTGATAATTTCAACGT 178  
 QY 736 CTCGAAAGCTACGAGATCGATGGCGGAGATCACAATGATGCTCTCCAAACCCCAAAA 795  
 Db 177 CTCGAAAGCTACGAGATCGATGGCGGAGATCACAATGATGCTCTCCAAACCCCAAAA 118  
 QY 796 ACTCTTTGACTCTCTCTCTGCTATTCGCCAGATATATGTAATATCTTAAGTCGGTTT 855  
 Db 117 ACTCTTTGACTCTCTCTCTGCTATTCGCCAGATATATGTAATATCTTAAGTCGGTTT 58  
 QY 856 TACTTTTTTCTCATCGTTACTTAATAAAACAAACCCCTTTTTCGGGGCAACTTTTCATC 912  
 Db 57 TACTTTTTTCTCATCGTTACTTAATAAAACAAACCCCTTTTTCGGGGCAACTTTTCATC 1  
 RESULT 7  
 AV824147  
 LOCUS AV824147 RAFL6 Arabidopsis thaliana CDNA clone RAFL06-14-F12 5',  
 DEFINITION AV824147.1 GI:19866207  
 mRNA sequence.  
 ACCESSION AV824147  
 VERSION AV824147.1  
 KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 441)  
 Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,  
 Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,  
 Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.  
 and Shinozaki, K.  
 Large scale analysis of Arabidopsis full-length cDNA (2002b)  
 Unpublished (2002)  
 Contact: Motoaki Seki  
 Plant Functional Genomics Research Group  
 RIKEN Genomic Sciences Center  
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
 Tel: 81-298-36-4359  
 Fax: 81-298-36-9060  
 Email: msek@rtc.riken.go.jp  
 An Arabidopsis full-length cDNA library was constructed essentially  
 as reported previously (Seki et al., 1998). This clone is in a  
 modified pBluescript vector as a SstI/XhoI insert. Please visit our  
 web site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for  
 further details.

TITLE  
 JOURNAL  
 COMMENT

CONTACT: Motoaki Seki  
 Plant Functional Genomics Research Group  
 RIKEN Genomic Sciences Center  
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
 Tel: 81-298-36-4359  
 Fax: 81-298-36-9060  
 Email: msek@rtc.riken.go.jp  
 An Arabidopsis full-length cDNA library was constructed essentially  
 as reported previously (Seki et al., 1998). This clone is in a  
 modified pBluescript vector as a SstI/XhoI insert. Please visit our  
 web site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for  
 further details.

further details.  
 Location/Qualifiers  
 1. .441  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="RAFL06-14-F12"  
 /dev\_stages="plants at various developmental stages from germination to mature seeds"  
 /lab\_host="DH10B"  
 /clone\_lib="RAFL6"  
 /note="Site 1: SstI; Site 2: XhoI; subjected to dehydration (1, 2, 5, 10, 24 hr) treatments"  
 47.7%; Score 434.8; DB 1; Length 441;  
 Best Local Similarity 99.1%; Pred. No. 1.4e-121;  
 Matches 436; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 ATCAAGCTTATCTCAACAAAGAAACCTTAACCAAAAGAAACCTCTACGGAGACA 60  
 Db 2 ATCAAGCTTATCTCAACAAAGAAACCTTAACCAAAAGAAACCTCTACGGAGACA 61  
 QY 61 TATGAGAGGAAACATCACTTCTGTTAGTTTCAACACGCTTATCATGGAGCTGGATCTG 120  
 Db 62 TATGAGAGGAAACATCACTTCTGTTAGTTTCAACACGCTTATCATGGAGCTGGATCTG 121  
 QY 121 GTACAAGCTCAAGCCCTCTTGAATCAGCGCGCCACCGCGTTACTGTGCGAACTCGC 180  
 Db 122 GTACAAGCTCAAGCCCTCTTGAATCAGCGCGCCACCGCGTTACTGTGCGAACTCGC 181  
 QY 181 CGCCTCCGGGATCGACCCACGACCAATCCAGCCGTTGAAACCGTCGACGAATCTCCAA 240  
 Db 182 CGCCTCCGGGATCGACCCACGACCAATCCAGCCGTTGAAACCGTCGACGAATCTCCAA 241  
 QY 241 ACCGTTGATCGAAACGCTCAATCTCTCCAGAGACGAGGTAATCTCGTTGGATT 300  
 Db 242 ACCGTTGATCGAAACGCTCAATCTCTCCAGAGACGAGGTAATCTCGTTGGATT 301  
 QY 301 CAGCTTCGGAGGCATCAACATCGCTCTCGCCCGGACATATTTCCGGCGAAGATTAAAGT 360  
 Db 302 CAGCTTCGGAGGCATCAACATCGCTCTCGCCCGGACATATTTCCGGCGAAGATTAAAGT 361  
 QY 361 TCTTGTGTTCTCAACGCTTCTTCCCGGACACCAACCGGCTTCTACGTTCTGGA 420  
 Db 362 TCTTGTGTTCTCAACGCTTCTTCCCGGACACCAACCGGCTTCTACGTTCTGGA 421  
 QY 421 CAAGTATATGGAGATGCCTG 440  
 Db 422 CAAGTTTCTGGANATGCCTG 441

RESULT 8  
 BP613009/c  
 LOCUS BP613009 RAFL16 Arabidopsis thaliana cdna clone RAFL16-91-K04 3',  
 DEFINITION mRNA sequence.  
 ACCESSION BP613009.1 GI:49264191  
 VERSION BP613009  
 KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 421)  
 Seki,M., Narusaka,M., Kamiya,A., Ishida,J., Satou,M., Sakurai,T.,  
 Nakajima,M., Enju,A., Akiyama,K., Oono,Y., Muramatsu,M.,  
 Hayashizaki,Y., Kawai,J., Carninci,P., Itoh,M., Ishii,Y.,  
 Arakawa,T., Shibata,K., Shinagawa,A. and Shinozaki,K.  
 Functional annotation of a full-length Arabidopsis cdna collection  
 Science 296 (5565), 141-145 (2002)

MEDLINE 21932900  
 PUBMED 11910074  
 COMMENT Contact: Motoaki Seki  
 Plant Functional Genomics Research Group  
 RIKEN Genomic Sciences Center  
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
 Tel: 81-298-36-4359  
 Fax: 81-298-36-9060  
 Email: meki@rtc.riken.go.jp  
 reversed clone; Please visit our web site  
 (http://pfigweb.gsc.riken.go.jp/) for further details.  
 FEATURES  
 Location/Qualifiers  
 1. .421  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="RAFL16-91-K04"  
 /lab\_host="DH10B"  
 /clone\_lib="RAFL16"  
 /note="Site 1: BamHI; Site 2: SalI; dark-grown"  
 ORIGIN  
 Query Match 46.2%; Score 421; DB 5; Length 421;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-117;  
 Matches 421; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 492 TGAGTTTATGAAGATGGGACCAAAATTCATGAAGGACGCTTTTACCAAAATTTGCCA 551  
 Db 421 TGAGTTTATGAAGATGGGACCAAAATTCATGAAGGACGCTTTTACCAAAATTTGCCA 362  
 QY 552 TAGAGGATTACGAGCTGCAAAAATGTTGCATAGCAAGGTCATTTTTCACAGGATC 611  
 Db 361 TAGAGGATTACGAGCTGCAAAAATGTTGCATAGCAAGGTCATTTTTCACAGGATC 302  
 QY 612 TATCAAAAGAAAGAAAGTTTAGCGAGGAAGATATGGTTCGGTGCAACGAGTTTACGTAA 671  
 Db 301 TATCAAAAGAAAGAAAGTTTAGCGAGGAAGATATGGTTCGGTGCAACGAGTTTACGTAA 242  
 QY 672 TGAGTAGTGAAGACAAAGCCATCCCTCGGATTTTCATTCGTTGGATGATGATTAATTCA 731  
 Db 241 TGAGTAGTGAAGACAAAGCCATCCCTCGGATTTTCATTCGTTGGATGATGATTAATTCA 182  
 QY 732 ACGTCTCGAAGTCTAGAGATCGATGCGGAGATCACATCGTGTCTTCCAAACCCC 791  
 Db 181 ACGTCTCGAAGTCTAGAGATCGATGCGGAGATCACATCGTGTCTTCCAAACCCC 122  
 QY 792 AAAAATCTTTTGACTCTCTCTCTATTTGCCACCGATTATATGTAATAATCTTAAAGTCC 851  
 Db 121 AAAAATCTTTTGACTCTCTCTCTATTTGCCACCGATTATATGTAATAATCTTAAAGTCC 62  
 QY 852 GTTTTACCTTTTCTCATCGTTACTATAAACAACCCCTTTTCCGGGCAACTTTCAT 911  
 Db 61 GTTTTACCTTTTCTCATCGTTACTATAAACAACCCCTTTTCCGGGCAACTTTCAT 2  
 QY 912 C 912  
 Db 1 C 1  
 RESULT 9  
 AA404853  
 LOCUS 28251 Lambda-PRL2 Arabidopsis thaliana cdna clone 143M11XP 3', mRNA  
 DEFINITION sequence.  
 ACCESSION AA404853  
 VERSION AA404853.1 GI:2062871  
 KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 480)

**AUTHORS** Newman, T., deBrujin, F. J., Green, P., Keestra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E. and Somerville, C.  
**TITLE** Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
**JOURNAL** Plant Physiol. 106, 1241-1255 (1994)  
**MEDLINE** 95148729  
**PUBMED** 7846151  
**COMMENT** Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22313tcrn@ibm.cl.msu.edu  
Seq primer: M13 -21 dye primer.  
Location/Qualifiers  
1. .480

**FEATURES**  
source  
1. .480  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/clone\_lib="143M11P"  
/clone\_lib="Lambda-PRL2"  
/note="Vector: lambda Zip-Lox; Site 1: Sal; Site 2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA."

**ORIGIN**  
Query Match 45.9%; Score 418.6; DB 1; Length 480;  
Best Local Similarity 98.4%; Pred. No. 1.3e-116;  
Matches 421; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 475 AACAGAAATGGACGATGTTTATGGAAGTGGACCAAAATTCATGAAGCGACGTCT 534  
DB 53 AACAGAAATGGACGATGTTTATGGAAGTGGACCAAAATTCATGAAGCGACGTCT 112  
QY 535 TTACCAAAATTCATGAGATGTTACGCTGGCAAAATGTTGCATAGCGAAGGTC 594  
DB 113 TTACCAAAATTCATGAGATGTTACGCTGGCAAAATGTTGCATAGCGAAGGTC 172  
QY 595 ATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTACGAGGAGGATGTTGTTTCGT 654  
DB 173 ATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTACGAGGAGGATGTTGTTTCGT 232  
QY 655 GCAACGAGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTCGATTCATTCGTTG 714  
DB 233 GCAACGAGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTCGATTCATTCGTTG 292  
QY 715 GATGATTGATATTTCAAGCTCTCGAAGTCTACGAGATCGATGGGAGATCACATGTT 774  
DB 293 GATGATTGATATTTCAAGCTCTCGAAGTCTACGAGATCGATGGGAGATCACATGTT 352  
QY 775 GATGCTCTCAAAACCCCAAAACTCTTTGACTCTCTCTGCTATTTGCCACCGATTATAT 834  
DB 353 GATGCTCTCAAAACCCCAAAACTCTTTGACTCTCTCTGCTATTTGCCACCGATTATAT 412  
QY 835 GTAATAATCTTAAGTCGGTTTTTACTTTTTTCTCATCGTTTACTTAATAAAACCAACCCCTTT 894  
DB 413 GTAATAATCTTAAGTCGGTTTTTACTTTTTTCTCATCGTTTACTTAATAAAACCAACCCCTTT 472  
QY 895 TTCCGGGC 902  
DB 473 TTCCGGGC 480

**RESULT** 10  
BP642196/c  
**LOCUS** BP642196  
**DEFINITION** mRNA sequence.  
**ACCESSION** BP642196  
**VERSION** BP642196  
**KEYWORDS** EST.  
**SOURCE** Arabidopsis thaliana (thale cress)  
**ORGANISM** Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

**REFERENCE**  
1 (bases 1 to 425)  
**AUTHORS** Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T., Nakajima, M., Enju, A., Akiyama, K., Oono, Y., Muramatsu, M., Hayashizaki, Y., Kawai, J., Carninci, P., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A. and Shinozaki, K.  
**TITLE** Functional annotation of a full-length Arabidopsis cDNA collection  
**JOURNAL** Science 296 (5565), 141-145 (2002)  
**MEDLINE** 21932900  
**PUBMED** 11910074  
**COMMENT** Contact: Motoaki Seki  
Plant Functional Genomics Research Group  
RIKEN Genomic Sciences Center  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-4359  
Fax: 81-298-36-9060  
Email: msekior@c.riken.go.jp  
reversed clone; Please visit our web site  
(http://pfweb.gsc.riken.go.jp/) for further details.

**FEATURES**  
source  
1. 425  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone="RAPL19-57-N06"  
/tissue\_type="mixture of silique and flower"  
/lab\_host="DH10B"  
/clone\_lib="RAPL19"  
/note="Site\_1: BamHI; Site\_2: SalI; Subtraction Library"

**ORIGIN**  
Query Match 43.6%; Score 397.4; DB 5; Length 425;  
Best Local Similarity 98.5%; Pred. No. 4.1e-110;  
Matches 401; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 506 ATGGACCAAAATTCATGAAGCGACGTCTTTACCAAAATTTCCCATAGAGGATTCAG 565  
DB 425 ATGGACCAAAATTCATGAAGCGACGTCTTTACCAAAATTTCCCATAGAGGATTCAG 366  
QY 566 CTGGCAAAATTTGCATAGCAAGGGTCATTTTTCACAGAGGATCTATCAAGAAAGAA 625  
DB 365 CTGGCAAAATTTGCATAGCAAGGGTCATTTTTCACAGAGGATCTATCAAGAAAGAA 306  
QY 626 AAGTTTACGAGGAGGATATGGTTTCGTCGCAACGAGTTTACGTAATCAGTAGTCAAGAC 685  
DB 305 AAGTTTACGAGGAGGATATATGGTTTCGTCGCAACGAGTTTACCTAATGAGTAGTCAAGAC 246  
QY 686 AAAGCCATCCCTCGCATTTTCATTCGTTGGATGATTAATTTCAACGTCCTCGAAAGTC 745  
DB 245 AAAGCCATCCCTCGCATTTTCATTCGTTGGATGATTAATTTCAACGTCCTCGAAAGTC 186  
QY 746 TACGAGATCGATGGCGGAGATCACATGGTATGCTCTCCAAACCCCAAAACTCTTTGAC 805  
DB 185 TACGAGATCGATGGCGGAGATCACATGGTATGCTCTCCAAACCCCAAAACTCTTTGAC 126  
QY 806 TCTCTCTGCTATTGCCACCGATTATATGTAATAATCTTAAGTCGCTTTTACTTTTTTC 865  
DB 125 TCTCTCTGCTATTGCCACCGATTATATGTAATAATCTTAAGTCGCTTTTACTTTTTTC 66  
QY 866 TCATCGTTTACTTAATAAAACCAACCCCTTTTCCGGCAACTTTTCATC 912

```

Db
|||||
65 TCTCGTTACTAATAAAACAACCCCTTTTCCGGCCAACTTTCATC 19

RESULT 11
AV790585/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 426)
Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
and Shinozaki,K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). This clone is in a
modified pBluescript vector as a SstI/XhoI insert. Please visit our
web site (http://www.gsc.riken.go.jp/e/plant/index\_e.html) for
further details.
FEATURES
source
1..426
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAF106-89-G15"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH10B"
/clone_lib="RAF16"
/note="Site 1: SstI; Site 2: XhoI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"
ORIGIN
Query Match 42.9%; Score 391.2; DB 1; Length 426;
Best Local Similarity 99.0%; Pred. No. 3.3e-108;
Matches 404; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 506 ATGGGACCAAAATTCATGAAGCAGCTCTTTACCAAAATTTGCCATAGAGGATTACGAG 565
Db 426 ATGGGACCAAAATTCATGAAGCAGCTCTTTACCAAAATTTGCCATAGAGGATTACGAG 367
QY 566 CTGGCAAAATCTTGTCATAGGCAAGGTCATTTTTCACAGAGGATCTATCAAGAAGAA 625
Db 366 CTGGCAAAATCTTGTCATAGGCAAGGTCATTTTTCACAGAGGATCTATCAAGAAGAA 307
QY 626 AAGTTTACGAGGAAGGATATCGTTTCGTTGCAACGAGTTTACGTAATGAGTAGTGAAGAC 685
Db 306 AAGTTTACGAGGAAGGATATTTGTTTCGTTGCAACGAGTTTACATATGAGTAGTGAAGAC 247
QY 686 AAGCCATCCCTGCGATTTCATTGTTGGATGA-TTGNATATTTCAACGTTCTCGAAGT 744
Db 246 AAGCCATCCCTGCGATTTCATTGTTGGATGAGTTGATATTTCAACGTTCTCGAAGT 187
QY 745 CTACGAGATCGATGGCGGAGATCATGTGTGATGCTCTCCAAACCCCAAAACTCTTTGA 804
|||||
186 GTACGAGATCGATGGCGGAGATCACATGGTGTGCTCTCCAAACCCCAAAACTCTTTGA 127
QY 805 CTCTCTCTCTGCTATTGCGACCGGATTATATGTAATAATCTTAAGTCGGTTTACTTTTT 864
Db 126 CTCTCTCTCTGCTATTGCGACCGGATTATATGTAATAATCTTAAGTCGGTTTACTTTTT 67
QY 865 CTATCGTTACTAATAAAACAACCCCTTTTTCGGGCAACTTTTCATC 912
Db 66 CTATCGTTACTAATAAAACAACCCCTTTTTCGGGCAACTTTTCATC 19

RESULT 12
AV820476/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 390)
Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
and Shinozaki,K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index\_e.html) for further
details.
FEATURES
source
1..390
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAF111-10-111"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH10B"
/clone_lib="RAF111"
/note="Site 1: BamHI; Site 2: SalI; subjected to various
treatments (dehydration, cold, high salt, ABA, heat and
UV). Dark-grown plants"
ORIGIN
Query Match 41.0%; Score 374.2; DB 1; Length 390;
Best Local Similarity 99.0%; Pred. No. 5.2e-103;
Matches 387; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 522 TGAAGGCACGCTTTTACCAAAATTTGCCATAGAGGATTACGAGTCGCAAAAATGTTGC 581
Db 390 TGAAGGCACGCTTTTACC-AAATGGTCCATAGAGGATTACGATCTGCGAAAATGTTGC 332
QY 582 ATAGCAAGGTCATTTTTCACAGAGGATCTCATCAAGAAAGAAAGCTTTTACCGGGAAG 641
Db 331 ATAGCAAGGTCATTTTTCACAGAGGATCTCATCAAGAAAGAAAGCTTTTACCGGGAAG 272
QY 642 GATATGGTTCGTCGACACGAGTTTACGTAATGAGTAGTGAAGACCAAGCCATCCCCTCG 701

```



Db 271 GATATGCTTCGGTGCACAGGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTGGG 212  
 QY 702 ATTTTCATTTCGTCGATGATTGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGG 761  
 Db 211 ATTTTCATTTCGTCGATGATTGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGGG 152  
 QY 762 GAGATCACATGCTGCTCTCCAAACCCCAAAACTCTTTGACTCTCTCTCTGCTATTG 821  
 Db 151 GAGATCACATGCTGCTCTCCAAACCCCAAAACTCTTTGACTCTCTCTCTGCTATTG 92  
 QY 822 CGACCGATTATGTTAATCTTAAGTCGGTTTACTTTTCTCATCGTTACTAATAA 881  
 Db 91 CCACCGATTATGTTAATCTTAAGTCGGTTTACTTTTCTCATCGTTACTAATAA 32  
 QY 882 AACAAACCCCTTTTCCGGGCAACTTTTCATC 912  
 Db 31 AACAAACCCCTTTTCCGGGCAACTTTTCATC 1

RESULT 13  
 AV554161  
 LOCUS 366 bp mRNA linear EST 23-FEB-2004  
 DEFINITION AV554161 Arabidopsis thaliana roots Columbia Arabidopsis thaliana  
 cDNA clone RZ82h03R 5', mRNA sequence.  
 ACCESSION AV554161  
 VERSION AV554161.1 GI:8725574  
 KEYWORDS EST  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE 1 (bases 1 to 366)  
 AUTHORS Asamizu,S., Nakamura,Y., Sato,S. and Tabata,S.  
 TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation  
 of 12,028 non-redundant expressed sequence tags from normalized and  
 size-selected cDNA libraries  
 JOURNAL DNA Res. 7 (3), 175-180 (2000)  
 MEDLINE 20363093  
 PUBMED 10907847  
 COMMENT Contact: Erika Asamizu  
 The First Laboratory for Plant Gene Research  
 Kazusa DNA Research Institute  
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
 Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

FEATURES  
 source  
 1..366  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /ecotype="Columbia"  
 /db\_xref="taxon:3702"  
 /clone="RZ82h03R"  
 /tissue\_type="roots"  
 /clone\_lib="Arabidopsis thaliana roots Columbia"  
 /note="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2:  
 XhoI"  
 ORIGIN  
 Query Match 40.1%; Score 366; DB 1; Length 366;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-100;  
 Matches 366; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 AACATCACTTCGTTAGTTTCACACGCTTATCATGGACCTGGATCTGTTCAAGCTCA 131  
 Db 1 AACATCACTTCGTTAGTTTCACACGCTTATCATGGACCTGGATCTGTTCAAGCTCA 60  
 QY 132 AGCCCCCTCTTGAATCAGCGGCGCACCGGTTACTGCTCGAACTCGCGGCTCCGGGA 191  
 Db 61 AGCCCCCTCTTGAATCAGCGGCGCACCGGTTACTGCTCGAACTCGCGGCTCCGGGA 120  
 QY 192 TCGACCCACGACCAATCCAGGCGGTTGAACCGTGCAGCAATACTCCAAACCGTTGATCG 251

Db 121 TCGACCCACGACCAATCCAGGCCGTTGAACCGTCGACGAATATCTCCAAACCGTTGATCG 180  
 QY 252 AAACCCCTCAAATCTCTTCAGAGAACGAAGAGTAATTTCTGTTTGGATTTCAGCTTCGGAG 311  
 Db 181 AAACCCCTCAAATCTCTTCAGAGAACGAAGAGTAATTTCTGTTTGGATTTCAGCTTCGGAG 240  
 QY 312 GCATCAACATCGCTCTCGCCGCCGACATATTTCCGGCCGAAGATTAAAGGTTCTTGTGTTCC 371  
 Db 241 GCATCAACATCGCTCTCGCCGCCGACATATTTCCGGCCGAAGATTAAAGGTTCTTGTGTTCC 300  
 QY 372 TCAACGCGCTTCTTCCGCCGACACCAACCGAGTGCCTTCTCACGTTCTGGAACAAGTATGCG 431  
 Db 301 TCAACGCGCTTCTTCCGCCGACACCAACCGAGTGCCTTCTCACGTTCTGGAACAAGTATGCG 360  
 QY 432 AGATGC 437  
 Db 361 AGATGC 366

RESULT 14  
 BP642613/c  
 LOCUS 405 bp mRNA linear EST 27-JUN-2004  
 DEFINITION BP642613 RAPL19 Arabidopsis thaliana cDNA clone RAPL19-59-D05 3',  
 mRNA sequence.  
 ACCESSION BP642613  
 VERSION BP642613.1 GI:49294083  
 KEYWORDS EST  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE 1 (bases 1 to 405)  
 AUTHORS Seki,M., Narusaka,M., Kamiya,A., Ishida,J., Satou,M., Sakurai,T.,  
 Nakajima,M., Enju,A., Akiyama,K., Oono,Y., Muramatsu,M.,  
 Hayashizaki,Y., Kawai,J., Carninci,P., Itoh,M., Ishii,Y.,  
 Arakawa,T., Shibata,K., Shinagawa,A. and Shinozaki,K.  
 TITLE Functional annotation of a full-length Arabidopsis cDNA collection  
 JOURNAL Science 296 (5565), 141-145 (2002)  
 MEDLINE 21932900  
 PUBMED 11910074  
 COMMENT Contact: Motoaki Seki  
 Plant Functional Genomics Research Group  
 RIKEN Genomic Sciences Center  
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
 Tel: 81-298-36-4359  
 Fax: 81-298-36-9060  
 Email: mseki@rtc.riken.go.jp  
 reversed clone; Please visit our web site  
 (http://pfgweb.gsc.riken.go.jp/) for further details.

FEATURES  
 source  
 1..405  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="RAPL19-59-D05"  
 /tissue\_type="mixture of silique and flower"  
 /lab\_host="DH10B"  
 /clone\_lib="RAPL19"  
 /note="Site\_1: BamHI; Site\_2: SalI; Subtraction Library"

ORIGIN  
 Query Match 39.5%; Score 359.8; DB 5; Length 405;  
 Best Local Similarity 99.0%; Pred. No. 1.4e-98;  
 Matches 383; Conservative 0; Mismatches 2; Indels 2; Gaps 2;  
 QY 528 CACGCTTTACCAAAATTTGTCCCATAGAGGATTTACGAGCTGGCAAAAATGTTGCATAGG 586  
 Db 405 CACGCTTTACCAAAATTTGTCCCATAGAGGATTTACGAGCTGGCAAAAATGTTGCATAGC 346  
 QY 587 CAAGGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTACGAGGAAGGATAT 646  
 Db 345 CAAGGGTCATTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTACGAGGAAGGATAT 286

|                       |     |  |     |
|-----------------------|-----|--|-----|
| QY                    | 647 | GTTTCGGTGCAACGAGTTTACCTAATAGTAGTAGTGAAGACAAAGCCATCCCCTGCGATTTC | 706 |
| Db                    | 285 | GTTTCGGTGCAACGAGTTTACCTAATAGTAGTAGTGAAGACAAAGCCATCCCCTGCGATTTC | 226 |
| QY                    | 707 | ATTTCGTTGGATGATGATTAATTTCAACGCTCGAAGTCTACGAGATCGATGGCGGAGAT    | 766 |
| Db                    | 225 | ATTTCGTTGGATGATGATTAATTTCAACGCTCGAAGTCTACGAGATCGATGGCGGAGAT    | 166 |
| QY                    | 767 | CACATGGTGATGCTCTCCAAACCCCAAAACCTCTTTGACTCTCTCTCTGCTA-TTGCCAC   | 825 |
| Db                    | 165 | CACATGGTGATGCTCTCCAAACCCCAAAACCTCTTTGACTCTCTCTCTGCTAGTTGCCAC   | 106 |
| QY                    | 826 | CGATTATATGTAATAATCTTAAGTCGTTTACTTTTTCATCGTTACTTAATAAACA        | 885 |
| Db                    | 105 | CGATTATATGTAATAATCTTAAGTCGTTTACTTTTTCATCGTTACTTAATAAACA        | 46  |
| QY                    | 886 | RAACCCCTTTTCCGGGCACTTTCATC                                     | 912 |
| Db                    | 45  | AAACCCCTTTTCCGGGCACTTTCATC                                     | 19  |
| RESULT 15             |     |  |     |
| AI993246              |     |  |     |
| LOCUS                 |     |  |     |
| DEFINITION            |     |  |     |
| AI993246              |     |  |     |
| ACCESSION             |     |  |     |
| VERSION               |     |  |     |
| KEYWORDS              |     |  |     |
| SOURCE                |     |  |     |
| ORGANISM              |     |  |     |
| REFERENCE             |     |  |     |
| AUTHORS               |     |  |     |
| TITLE                 |     |  |     |
| JOURNAL               |     |  |     |
| COMMENT               |     |  |     |
| FEATURES              |     |  |     |
| Source                |     |  |     |
| ORIGIN                |     |  |     |
| Query Match           |     |  |     |
| Best Local Similarity |     |  |     |
| Matches               |     |  |     |
| QY                    |     |  |     |
| Db                    |     |  |     |
| QY                    |     |  |     |

|    |     |   |     |
|----|-----|---|-----|
| Db | 61  | CATCACTTCGTGTAGTTTACAAACGCTTATCATCGAGCCTGGATCTGGTACAAGCTCAAG  | 120 |
| QY | 134 | CCCCCTCCTTGAATCAGCCGCGCCACCGGTTACTGCTGTGAACTCGCGCGCTCCGGGATC  | 193 |
| Db | 121 | CCCCCTCCTTGAATCAGCCGCGCCACCGGTTACTGCTGTGAACTCGAGCGCTCCGGGATC  | 180 |
| QY | 194 | GACCCAGACCAATCCAGGCGGTTGAAACCGTCGACGAATACTCCAAACCGTTGATCGAA   | 253 |
| Db | 181 | GACCCAGACCAATCCAGGCGGTTGAAACCGTCGACGAATACTCCAAACCGTTGATCGAA   | 240 |
| QY | 254 | ACCTCTCAAATCTCTTCCAGAGAACGAAGAGTAACTCTGGTTGGATTTCAGCTTCGGAGGC | 313 |
| Db | 241 | ACCTCTCAAATCTCTTCCAGAGAACGAAGAGTAACTCTGGTTGGATTTCAGCTTCGGAGGC | 300 |
| QY | 314 | ATCAACATCGCTCTCGCGCGGACATATTTCCGCGGAAGATTAAAGTTCTTTGTTCCTC    | 373 |
| Db | 301 | ATCAACATCGCTCTCGCGCGGACATATTTCCGCGGAAGATTAAAGTTCTTTGTTCCTC    | 360 |
| QY | 374 | A   | 374 |
| Db | 361 | A   | 361 |

Search completed: September 12, 2005, 06:31:55  
Job time : 2912.77 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: September 11, 2005, 22:50:50 ; Search time 499.745 Seconds  
(without alignments)  
10803.121 Million cell updates/sec

Title: US-10-780-002-36

Perfect score: 912

Sequence: 1 atcagctattctcaacaa.....ttttcgggcaactttcatc 912

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: geneseqn1980s:\*
- 2: geneseqn1990s:\*
- 3: geneseqn2000s:\*
- 4: geneseqn2001as:\*
- 5: geneseqn2001bs:\*
- 6: geneseqn2002as:\*
- 7: geneseqn2002bs:\*
- 8: geneseqn2003as:\*
- 9: geneseqn2003bs:\*
- 10: geneseqn2003cs:\*
- 11: geneseqn2003ds:\*
- 12: geneseqn2004as:\*
- 13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Match | Length | DB ID | Description        |
|------------|-------|-------|--------|-------|--------------------|
| 1          | 912   | 100.0 | 984    | 8     | ADA26534 Arabidops |
| 2          | 910.4 | 99.8  | 1056   | 3     | AAC47895 Arabidops |
| 3          | 887.8 | 97.3  | 909    | 3     | AAC47800 Arabidops |
| 4          | 481.4 | 52.8  | 483    | 10    | ADE82083 Arabidops |
| 5          | 243   | 26.6  | 918    | 8     | ADA26530 Arabidops |
| 6          | 241.6 | 26.5  | 792    | 6     | ABZ13515 Arabidops |
| 7          | 241.6 | 26.5  | 816    | 3     | AAC42344 Arabidops |
| 8          | 232.6 | 25.5  | 1079   | 8     | ADA26498 Arabidops |
| 9          | 222.6 | 24.4  | 1152   | 8     | ADA26532 Arabidops |
| 10         | 221.6 | 24.3  | 1091   | 2     | AAT36351 Hevea bra |
| 11         | 219.2 | 24.0  | 817    | 2     | AAT86824 Hevea bra |
| 12         | 218.8 | 24.0  | 792    | 3     | AAC43133 Arabidops |
| 13         | 218.8 | 24.0  | 792    | 6     | ABZ12412 Arabidops |
| 14         | 218.8 | 24.0  | 792    | 12    | ADN74620 Thale cre |
| 15         | 214.2 | 23.5  | 792    | 6     | ABZ13746 Arabidops |
| 16         | 214.2 | 23.5  | 864    | 3     | AAC47519 Arabidops |
| 17         | 214.2 | 23.5  | 870    | 3     | AAC34411 Arabidops |
| 18         | 214.2 | 23.5  | 950    | 8     | ADA26531 Arabidops |
| 19         | 205.4 | 22.5  | 777    | 12    | ADQ16305 Nucleotid |
| 20         | 201.8 | 22.1  | 823    | 3     | AAC42529 Arabidops |

|    |       |      |      |    |                    |
|----|-------|------|------|----|--------------------|
| 21 | 201.8 | 22.1 | 928  | 8  | ADA26538 Arabidops |
| 22 | 199.6 | 21.9 | 1109 | 8  | ADA26533 Arabidops |
| 23 | 198.6 | 21.8 | 777  | 12 | ADQ16306 Nucleotid |
| 24 | 194.4 | 21.3 | 771  | 6  | ABZ13473 Arabidops |
| 25 | 193.4 | 21.2 | 770  | 8  | ADA67827 Arabidops |
| 26 | 192.2 | 21.1 | 834  | 8  | ADA26537 Arabidops |
| 27 | 185.2 | 20.3 | 783  | 8  | ADA26536 Arabidops |
| 28 | 178.6 | 19.6 | 795  | 8  | ADA70991 Rice gene |
| 29 | 176.8 | 19.4 | 1046 | 8  | ADA26535 Arabidops |
| 30 | 172.6 | 18.9 | 828  | 3  | AAC41733 Arabidops |
| 31 | 172.6 | 18.9 | 903  | 3  | AAC34141 Arabidops |
| 32 | 172.6 | 18.9 | 908  | 8  | ADA26539 Arabidops |
| 33 | 141.2 | 15.5 | 422  | 3  | AAC43356 Arabidops |
| 34 | 136.2 | 14.9 | 984  | 3  | AAC49239 Arabidops |
| 35 | 122.6 | 13.4 | 1325 | 3  | AAC47916 Arabidops |
| 36 | 115.6 | 12.7 | 783  | 8  | ADA70766 Rice gene |
| 37 | 96.6  | 10.6 | 978  | 8  | ADA69955 Rice gene |
| 38 | 91.4  | 10.0 | 860  | 8  | ADA71305 Rice gene |
| 39 | 91.2  | 10.0 | 1041 | 8  | ADA26546 Arabidops |
| C  |       |      |      |    |                    |
| 40 | 91    | 10.0 | 541  | 10 | ABX57306 Arabidops |
| 41 | 72.4  | 7.9  | 910  | 3  | AAC35440 Arabidops |
| 42 | 72.4  | 7.9  | 1516 | 8  | ADA26547 Arabidops |
| 43 | 71.4  | 7.8  | 1815 | 3  | AAC44497 Rice maye |
| 44 | 71.2  | 7.8  | 720  | 8  | ADA70729 Rice gene |
| 45 | 68.4  | 7.5  | 789  | 6  | ABZ12505 Arabidops |

## ALIGNMENTS

RESULT 1  
ADA26534  
ID ADA26534 standard; DNA; 984 BP.  
XX  
AC ADA26534;

DT 20-NOV-2003 (first entry)

DE Arabidopsis thaliana AtSB2L5 gene homologous to tobacco SABP2 gene.  
XX  
KW salicylic acid-binding protein; SABP2; disease resistance; plant;  
modulator; tobacco; ds; gene.  
XX  
OS Arabidopsis thaliana.

PN WO2003016551-A2.

PD 27-FEB-2003.

PF 16-AUG-2002; 2002WO-US026312.

PR 16-AUG-2001; 2001US-0312863P.

XX (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

PI Klessig DF, Kumar D;

DR WPI; 2003-278578/27.

PT New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for  
producing transgenic plants with increased resistance to disease, or for  
screening SABP2 modulators that confer enhanced resistance of plants to  
disease.

PS Claim 65; Page; 98pp; English.

CC The invention relates to a novel isolated 1079 bp salicylic acid-binding  
protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid  
protein, sequences that specifically hybridize with the nucleic acid, the  
complement of the nucleic acid or a natural allelic variant of the  
sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for  
conferring increased resistance to disease in plants, or for screening  
modulators of SABP2, which confer increased or enhanced resistance of



PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0141843P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.

PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158023P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159323P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 99.8%; Score 910.4; DB 3; Length 1056;

Best Local Similarity 99.9%; Pred. No. 6.8e-286;

Matches 911; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCAGCTATTCTCAACAAAGAACCTTTAACCAAAAAGAAACGTTTACGGAGACA 60

Db 128 ATCAGCTATTCTCAACAAAGAACCTTTAACCAAAAAGAAACGTTTACGGAGACA 187

QY 61 TATGAGAGGAACATCACTCGTGTAGTTACACACGCTTATCATGAGCGCTGGATCTG 120  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 188 TATGAGAGGAACATCACTCGTGTAGTTACACACGCTTATCATGAGCGCTGGATCTG 247  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 121 GTAACAAGCTCAAGCCGCTCTTGAATCAGCCGGCCACCGGTTACTGCTGCGAACTCGC 180  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 248 GTAACAAGCTCAAGCCGCTCTTGAATCAGCCGGCCACCGGTTACTGCTGCGAACTCGC 307  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 181 CGCCTCCGGGATCGACCCACGACCAATCCAGGCGCTTGAACCGTCGACGAATATCCAA 240  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 308 CGCCTCCGGGATCGACCCACGACCAATCCAGGCGCTTGAACCGTCGACGAATATCCAA 367  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 241 ACCGTTGATCGAAGCCCTCAATCTCTCCAGAGAACGAGAGGTAATCTCGTTGGAAT 300  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 368 ACCGTTGATCGAAGCCCTCAATCTCTCCAGAGAACGAGAGGTAATCTCGTTGGAAT 427  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 301 CAGCTTCGGAGGCATCAACATCGCTCTCGCCGCGACATATTTCCGGCGAAGATTAAGGT 360  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 428 CAGCTTCGGAGGCATCAACATCGCTCTCGCCGCGACATATTTCCGGCGAAGATTAAGGT 487  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 361 TCTTGTGTTCTCAACGCTTCTTCCCGGACACAAACCGTCTTCTACGTTCTGGA 420  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 488 TCTTGTGTTCTCAACGCTTCTTCCCGGACACAAACCGTCTTCTACGTTCTGGA 547  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 421 CAGCTTCGGAGGCATCAACATCGCTCTCGCCGCGACATATTTCCGGCGAAGATTAAGGT 480  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 548 CAGCTTCGGAGGCATCAACATCGCTCTCGCCGCGACATATTTCCGGCGAAGATTAAGGT 607  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 481 AAATGGGACGATGAGTTTATCAAGATGGGACCAAAATTCATGAAGGCACGCTTTTACCA 540  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 608 AAATGGGACGATGAGTTTATCAAGATGGGACCAAAATTCATGAAGGCACGCTTTTACCA 667  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 541 AAATGTCCTCCATGAGGATTCAGAGCTGGCAAAATGTTGCATAGGCAAGGTCATTTT 600  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 668 AAATGTCCTCCATGAGGATTCAGAGCTGGCAAAATGTTGCATAGGCAAGGTCATTTT 727  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 601 CACAGAGGATCTATCAAGAAAGAAAGTTAGCGAGGAGGATATGTTCCGGTCCACG 660  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 728 CACAGAGGATCTATCAAGAAAGAAAGTTAGCGAGGAGGATATGTTCCGGTCCACG 787  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 661 AGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTCGGATTTCAATTCGTTGGATGAT 720  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 788 AGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTCGGATTTCAATTCGTTGGATGAT 847  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 721 TGATAATTTCAACGCTCGAAAGTCTACGAGATCGATGCGGAGATCAATGGTCAATGCT 780  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 848 TGATAATTTCAACGCTCGAAAGTCTACGAGATCGATGCGGAGATCAATGGTCAATGCT 907  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 781 CTCCAAACCCCAAACTCTTTGACTCTCTCTGCTATTCGACCGGATTAATGTAATA 840  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 908 CTCCAAACCCCAAACTCTTTGACTCTCTCTGCTATTCGACCGGATTAATGTAATA 967  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 841 ATCTTAAGTCGCTTTTACTTTTCTCATCGTTACTTAATAAAACCAACCCCTTTTCCGG 900  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 968 ATCTTAAGTCGCTTTTACTTTTCTCATCGTTACTTAATAAAACCAACCCCTTTTCCGG 1027  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 901 GCAACTTTCATC 912  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1028 GCAACTTTCATC 1039  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 3

AAC47800  
ID AAC47800 standard; DNA; 909 BP.  
XX  
AC AAC47800;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 55168.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;

KW promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
OS EP1033405-A2.  
XX  
PN 06-SEP-2000.  
PD  
PF 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
XX 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126284P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.

```
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-014342P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 26-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.

PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 97.3%; Score 887.8; DB 3; Length 909;
Beat Local Similarity 99.7%; Pred. No. 1.5e-278;
Matches 900; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATCAGCTATTCTCAACAAGCAAGAACCTTAAACCAAAAAGAAACCTTACCGGAGACA 60
DB 7 ATAAGCTATTCTCAACAAGCAAGAACCTTAAACCAAAAAGAAACCTTACCGGAGACA 66

QY 61 TATGAGAGAGAAACATCACTTCGTGTAGTTTCAACAACGCTTATCATGGAGCCTGGATCTG 120
DB 67 TATGAGAGAGAAACATCACTTCGTGTAGTTTCAACAACGCTTATCATGGAGCCTGGATCTG 126

QY 121 GTAAGAAGCTCAAGCCCTCTTGAATACAGCGGCCACCGCGTTACTGCTGCGAACTCGC 180
DB 127 GTAAGAAGCTCAAGCCCTCTTGAATACAGCGGCCACCGCGTTACTGCTGCGAACTCGC 186

QY 181 CGCCTCCGGATCGACCCACCAATCCAGCCGTTGAACCGTCGACGAATACTCCAA 240
DB 187 CGCCTCCGGATCGACCCACCAATCCAGCCGTTGAACCGTCGACGAATACTCCAA 246

QY 241 ACCGTTGATCGAAACCCCTCAAAATCTTCCAGAGAACGAAAGGTAATTTCTGGTTGGATT 300
DB 247 ACCGTTGATCGAAACCCCTCAAAATCTTCCAGAGAACGAAAGGTAATTTCTGGTTGGATT 306

QY 301 CAGCTTCGGAGGATCAACATCGCTCTCGCCCGGACATATTTCCGGCGAAGATTAAAGT 360
DB 307 CAGCTTCGGAGGATCAACATCGCTCTCGCCCGGACATATTTCCGGCGAAGATTAAAGT 366
```

```
QY 361 TCTTGTGTTCTCAAGCGCTTCTTGGCCGACACCAACCCACGTCCTTCTCAGTTCGGA 420
Db 367 TCTTGTGTTCTCAAGCGCTTCTTGGCCGACACCAACCCACGTCCTTCTCAGTTCGGA 426
QY 421 CAAGTATATGGAGATGCTTGGAGGTTTGGAGATTGTGAGTTTTCATCTCATGAAACAAG 480
Db 427 CAAGTATATGGAGATGCTTGGAGGTTTGGAGATTGTGAGTTTTCATCTCATGAAACAAG 486
QY 481 AAATGGGACGATGAGTATTTATGAAGATGGGACCAAAATTCATGAAGGACGCTTTTACCA 540
Db 487 AAATGGGACGATGAGTATTTATGAAGATGGGACCAAAATTCATGAAGGACGCTTTTACCA 546
QY 541 AAATGTCCTCATAGAGATTACGAGCTGGCAAAATGTTGCATAGCGAAGGTCATTTTT 600
Db 547 AAATGTCCTCATAGAGATTACGAGCTGGCAAAATGTTGCATAGCGAAGGTCATTTTT 606
QY 601 CACAGAGGATCTATCAAGAAAGAAAAGTTTAGCGAGGAAGGATATGTTGGTGCACAG 660
Db 607 CACAGAGGATCTATCAAGAAAGAAAAGTTTAGCGAGGAAGGATATGTTGGTGCACAG 666
QY 661 AGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTGCGATTTTCATTCGTTGGATGAT 720
Db 667 AGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTGCGATTTTCATTCGTTGGATGAT 726
QY 721 TGATTAATTTCAACGCTCTGAAAGTCTACGAGATCGATGCGGAGATCACATGGTGATGCT 780
Db 727 TGATTAATTTCAACGCTCTGAAAGTCTACGAGATCGATGCGGAGATCACATGGTGATGCT 786
QY 781 CTCCTAAACCCAAAACCTTTTGACTCTCTCTCTGCTATTTGCGCCGATTTATATGTAATA 840
Db 787 CTCCTAAACCCAAAACCTTTTGACTCTCTCTCTGCTATTTGCGCCGATTTATATGTAATA 846
QY 841 ATCTTT-AAAGTCGGTTTACTTTTTTCTCATCGTTACTAATAAAACCAACCCCTTTTTCGG 899
Db 847 ATCTTAAAGTCGTTTACTTTTTTCTCATCGTTACTAATAAAACCAACCCCTTTTTCGG 906
QY 900 GGC 902
Db 907 GGC 909

RESULT 4
ADE82083
ID ADE82083 standard; cDNA; 483 BP.
AC ADE82083;
XX
XX
DT 29-JAN-2004 (first entry)
DE
DE Arabidopsis thaliana expressed polynucleotide seq id 854.
XX
KW genetically modified organism; transgenic organism; plant;
KW inhibitor testing; activator testing; modifier testing; fungicide;
KW insecticide; genetic function; genetic regulation; cellular metabolism;
KW gene; ss.
XX
XX Arabidopsis thaliana.
XX
XX US2003115639-A1.
XX
XX 19-JUN-2003.
XX
XX 26-JAN-2001; 2001US-00770961.
XX
XX 27-JAN-2000; 2000US-0178466P.
XX
XX (GORLACH J.
XX (ANYI Y.
XX (HAMILTON C M.
XX (PRICE J L.
XX (RAIN/ RAINES T M.
XX (YUYI/ YU Y.
XX (RAME/ RAMEKA J G.
```

```
PA (PAGE/ PAGE A.
PA (MATH/ MATHW A V.
PA (LEDF/ LEDFORD B L.
PA (WOES/ WOESSNER J P.
PA (HAAS/ HAAS W D.
PA (GARC/ GARCIA C A.
PA (KRIC/ KRICKER M.
PA (SLAT/ SLATER T.
PA (DAVI/ DAVIS K R.
PA (ALLE/ ALLEN K.
PA (HOFF/ HOFFMAN N.
PA (HURB/ HURBAN P.
XX GORLACH J, AN Y, HAMILTON CM, PRICE JL, RAINES TM, YU Y;
PI RAMEKA JG, PAGE A, MATHW AV, LEDFORD BL, WOESSNER JP, HAAS WD;
PI GARCIA CA, KRICKER M, SLATER T, DAVIS KR, ALLEN K, HOFFMAN N;
PI HURBAN P;
XX WPI; 2003-810930/76.
XX
XX Novel Arabidopsis thaliana nucleic acids useful for generating
XX genetically modified transgenic organisms, for screening biologically
XX active agents such as fungicides, insecticides.
XX
XX Claim 1; SEQ ID NO 854; 44pp; English.
XX
XX The invention describes a nucleic acid (I) comprising a sequence capable
XX of hybridising under stringent conditions to any one of 999 fully defined
XX Arabidopsis thaliana sequences (I) as given in specification e.g., 360,
XX 1137, 455, 219, 472 nucleotides, etc, or its fragment. (I) is useful as a
XX hybridisation probe to complementary molecules in a cDNA library. (I) is
XX also useful for generating genetically modified and transgenic organisms,
XX usually plant cells and plants. A protein encoded by (I) is useful in
XX screening assays to determine the effect of candidate inhibitors,
XX activators or modifiers of the gene product. The protein is also useful
XX for screening biologically active agents e.g., fungicides and
XX insecticides. A genetically modified cell, comprising an exogenous
XX nucleic acid, where the nucleic acid comprises transcription regulatory
XX sequences operably linked to a sequence capable of hybridising under
XX stringent conditions to (I) is useful in the study of genetic function
XX and regulation, for alteration of the cellular metabolism and for
XX screening compounds that may affect the biological function of the gene
XX or gene product. This sequence represents an Arabidopsis thaliana
XX polynucleotide of the invention.
XX
XX Sequence 483 BP; 129 A; 127 C; 112 G; 115 T; 0 U; 0 Other;
SQ
Query Match 52.8%; Score 481.4; DB 10; Length 483;
Best Local Similarity 99.8%; Pred. No. 5.3e-146;
Matches 482; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 26 GAACCTTAAACCAAAAGAAAACGTTACGGAGACATATGGAGAGAAACATCATTCTG 85
Db 1 GAACCTTAAACCAAAAGAAAACGTTACGGAGACATATGGAGAGAAACATCATTCTG 60
QY 86 TTAGTTCAACACGCTTATCATGGAGCCTGGATCTGGTACAAGCTCAAGCCCTCTTGAA 145
Db 61 TTAGTTCAACACGCTTATCATGGAGCCTGGATCTGGTACAAGCTCAAGCCCTCTTGAA 120
QY 146 TCAGCCGCCACCGGTTACTGCTGAACTCCCGCTCCGGATCGACCCAGACCA 205
Db 121 TCAGCCGCCACCGGTTACTGCTGAACTCCCGCTCCGGATCGACCCAGACCA 180
QY 206 ATCCAGGCGGTTGAAACCGTCGACGAATACCTCCAAACCGTTGATCGAAATCT 265
Db 181 ATCCAGGCGGTTGAAACCGTCGACGAATACCTCCAAACCGTTGATCGAAATCT 240
QY 266 CTTCCAGAGAACGAAGGTAATTCGTTGGATTTCAGCTTCGAGGATCAACATCGCT 325
Db 241 CTTTCAGAGAACGAAGGTAATTCGTTGGATTTCAGCTTCGAGGATCAACATCGCT 300
QY 326 CTCGCGCGGACATATTTCCGCGCAAGATTAAAGTTCTTGTTCTCTCAACGCTTCTTG 385
```

Db 301 CTCGCCGCGACATATTTCCGGCGAAGATTAAAGTTCTTGTTCTCAACGCTTCTTG 360  
 QY 386 CCGGACACCAACCGCTCTCTACAGTTCTCGACAAAGTATATGAGATGCTCGGAGGT 445  
 Db 361 CCGGACACCAACCGCTCTCTACAGTTCTCGACAAAGTATATGAGATGCTCGGAGGT 420  
 QY 446 TTGGGAGATTGAGTTTTCATCTCATGAACAAAGAAATGGGACGATGAGTTTATTGAAG 505  
 Db 421 TTGGGAGATTGAGTTTTCATCTCATGAACAAAGAAATGGGACGATGAGTTTATTGAAG 480  
 QY 506 ATG 508  
 Db 481 ATG 483

## RESULT 5

ADA26530

ID ADA26530 standard; DNA; 918 BP.

XX AC

ADA26530;

XX 20-NOV-2003 (first entry)

XX Arabidopsis thaliana AtS2L1 gene homologous to tobacco SABP2 gene.

DE Arabidopsis thaliana

XX salicylic acid-binding protein; SABP2; disease resistance; plant;

KW modulator; tobacco; ds; gene.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

XX WO2003016551-A2.

XX 27-FEB-2003.

XX 16-AUG-2002; 2002WO-US026312.

XX 16-AUG-2001; 2001US-0312863P.

XX (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

XX Klessig DF, Kumar D;

XX WPI; 2003-278578/27.

XX New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for

PT producing transgenic plants with increased resistance to disease, or for

PT screening SABP2 modulators that confer enhanced resistance of plants to

PT disease.

XX Claim 65; Page: 98pp; English.

XX The invention relates to a novel isolated 1079 bp salicylic acid-binding

CC protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid

CC protein, sequences that specifically hybridize with the nucleic acid, the

CC complement of the nucleic acid or a natural allelic variant of the

CC sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for

CC conferring increased resistance to disease in plants, or for screening

CC modulators of SABP2, which confer increased or enhanced resistance of

CC plants to disease. The SABP2 or SABP2 homologue nucleic acid is also

CC useful for producing transgenic plants with increased resistance to

CC disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful

CC for studying the molecular mechanisms responsible for the modulation of

CC salicylic acid-mediated disease resistance in plants. This sequence

CC represents the gene encoding the Arabidopsis thaliana SABP2 AtS2L1 which

CC has homology to the gene (ADA26498) encoding the tobacco SABP2 protein.

CC (Note: this sequence is not given in the specification but is derived

CC from the Genbank accession number given by the inventors).

XX Sequence 918 BP; 266 A; 175 C; 221 G; 256 T; 0 U; 0 Other;

SQ

Query Match

Best Local Similarity

Matches 468; Conservative

26.6%; Score 243; DB 8; Length 918;

59.6%; Pred. No. 5.6e-68;

0; Mismatches 305; Indels 12; Gaps 3;

QY 66 AGAGGAAACATCACCTTCGTGTAGTTTACAAACGCTTATCATGAGCCTGGATCTGTACA 125  
 Db 57 AGAGGAAACACATTTTGTACTAGTACATGGTTCGTGCCATGCGGGTGTCTGTGTACA 116  
 QY 126 AGCTCAAGCCCTCCTTGAATCAGCCGCGCCACCGGTTTACTGCTGTGAACTCTGCGGCT 185  
 Db 117 AGGTTAAGCCGTGCTAGAGGCGGTGGGCCACCGCGTAACTGCTGTGGACTTGTGCTCCT 176  
 QY 186 CCGGATCGACCCCA---CGACCAATCCAGGCGGTGAAACCCGTCGACGATCTCAAC 242  
 Db 177 CCGGAATAGACACAAACGAGGTGATCACTGACATCCCCACATGCGCAACAATACTCGGAGC 236  
 QY 243 CGTTGATCGAAACCCCTCAAACTCTTCCAGAGAAAGAGAGTAATCTTGGTTGATTTCA 302  
 Db 237 CATTGACGAGCTCTGACCTCATTTGCCAATGATGAAAGGTTGCTCGTTGTGTACA 296  
 QY 303 GCTTCGGAGGCATCAACATCGCTTCGCGCGCGGACATATTTCCGCGCGAAGATTAAAGTTTC 362  
 Db 297 GCTTTGGTGGCTTGAACCTTAGCCATAGCCATGGAAGTTTCCCGGAAAAAATCTCTGTGCG 356  
 QY 363 TTGTTCTCTCAAGCCTTCTTGGCCGACACAAACCCAGCTGCTTCTCACGTTCTGGACA 422  
 Db 357 CTGTATTCTTGACTGCTTTCATGCGCGACACCGAAACACTCACCATCTCTTCGTTTGGACA 416  
 QY 423 AGTAT-----ATGGAGATGCTGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAA 476  
 Db 417 AGTTTGGAGCAACATGCTCAAGNAGCATGATGGGCAACCGAATTCGAACCTTATGTTT 476  
 QY 477 CAAGAAATGGGACGATGAGTTTATTGAAGATGGGACCAAAATTTCAATGAAGGCACTCTTTT 536  
 Db 477 CAGACAAATTCGGGACTGAGTATGTT---TTTAGCCCTGACTTTCATGAAGTTGGGTCTCT 533  
 QY 537 ACCAAATTTCCCATAGAGATTACGAGCTGGGCAAAATTTGTCATAGGCAAGGTCAT 596  
 Db 534 ACCAGCTTTTCCAGTTGAGGATCTTGAACCTGGGATTACTTTTAATGAGGCCAGGATCGT 593  
 QY 597 TTTTCACAGAGATCTATCAAGAAAGAAAGTTTATAGCGAGGAGGATATGTTTCGGTGC 656  
 Db 594 TATTATTAAAGTTTATCGAAGATGAAACTTCTCGGNTGAAAGATATGGGTCTGTTC 653  
 QY 657 AACGAGTTTACGTAAATGATAGTGAAGACAAAGCCATCCCTGCGAATTTCAATCGTTGA 716  
 Db 654 CTCGAGTTTTCATAGTGTGTAAGAGGACAAAGCAATTCAGAGAGAACGCCAGAGATGA 713  
 QY 717 TGATTGATAATTTCAACGCTTCGAAAGTCTAGAGATCGATGGCGGAGATCATCATGTTGA 776  
 Db 714 TGATTGATAATTTTCCCGTGAATTTAGTGTGAGATGGAGAGACAGATCATATATGCCAA 773  
 QY 777 TGCTCTCCAAACCCCAAAACTCTTTGACTCTCTCTGCTTATTCGCCACCGATTATATGT 836  
 Db 774 TGTTCGCAAGCTCAGCAACTCAGTGATTACTTCTGAAATTCGGGACAAATTCGTTT 833  
 QY 837 AATAA 841  
 Db 834 AATCA 838

## RESULT 6

ABZ13515

ID ABZ13515 standard; DNA; 792 BP.

XX AC

ABZ13515;

XX 21-JAN-2003 (first entry)

XX Arabidopsis thaliana stress regulated gene SEQ ID NO 1320.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX Arabidopsis thaliana.

XX WO200216655-A2.

XX 28-FEB-2002.  
PD  
XX PF  
XX 24-AUG-2001; 2001WO-US026685.  
XX  
XX 24-AUG-2000; 2000US-0227866P.  
PR 26-JAN-2001; 2001US-0264647P.  
PR 22-JUN-2001; 2001US-0300111P.  
XX  
XX (SCRI ) SCRIPPS RES INST.  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
XX Harper JF, Kreps J, Wang X, Zhu T;  
PI WPI; 2002-304127/34.  
DR  
XX  
XX Identifying a stress condition to which a plant cell has been exposed and  
PT producing plants with increased tolerance to these abiotic stresses.  
XX  
XX Claim 144; SEQ ID NO 1320; 577pp + Sequence Listing; English.  
XX  
XX The invention relates to identifying a stress condition to which a plant  
CC cell has been exposed, comprising: (a) contacting nucleic acid  
CC representative of expressed polynucleotides in the plant cell with an  
CC array or probes representative of the plant cell genome; and (b)  
CC detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present sequence is that  
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
CC in methods of the invention. Note: The sequence data for this patent is  
CC not represented in the printed specification but is based on sequence  
CC information supplied to Derwent by the European Patent Office  
XX  
SQ Sequence 792 BP; 217 A; 167 C; 196 G; 212 T; 0 U; 0 Other;  
Query Match 26.5%; Score 241.6; DB 6; Length 792;  
Best Local Similarity 59.6%; Pred. No. 1.5e-67;  
Matches 466; Conservative 0; Mismatches 304; Indels 12; Gaps 3;  
QY 66 AGAGGAAACATCATTGCTGTAGTTTACACACGCTTATCATGAGCGTGGATCTGTGACA 125  
DB 14 AGAGGAAACAAATTTGCTAGTATGATGATGCTGTGCTGCTGCTGCTGCTGCTGCTG 73  
QY 126 AGCTCAAGCCCTCTTGAATCAGCGCGCCACCGGTTACTGCTGTGCAACTCGCGGCT 185  
DB 74 AGGTTAAGCGGCTGTAGAGCGGTGGCCACCGGTAATGCTGTGCACTTAGCTGCT 133  
QY 186 CCGGGATCGACCA---CGACCAATCCAGCGGTTGAAACCGGTGACGAACTACTCAAAC 242  
DB 134 CCGGAATAGACACACAGGTCGATCACTGACATCCCAATCCCAATGCGAACAATACTCGGAGC 193  
QY 243 CGTTGATCGAAACCTCAATCTCTCCAGAGAACGAGAGCTAATCTGTTGGATTCA 302  
DB 194 CATTGCGAAGCTCTGACCTCATTCGCAATGATGAAAGGTGTGCTGTTGTGTCACA 253  
QY 303 GCTTCGAGGAGCTCAACATCGCTCTCGCGCGGACATATTTCCGGCGAAGATTAAAGTTTC 362  
DB 254 GCTTTGTTGCTTGAATCTTAGCCATAGCCATGAAAGTTTCCGAAATAATCTGTGCG 313  
QY 363 TTGTGTTCTCAAGCGCTTCTTGCCGCAACACCCAGCTGCTTCTACGTTCTTGAGCA 422  
DB 314 CTGTATTCTTGTGCTGCTTTCATGCGGCAACCGCAACATCACCATCTCTTCTGTTGGACA 373  
QY 423 AGTAT-----ATGGAGATGCTCGAGGTTTCGGAGTTGTGAGTTTCTCATGAAA 476  
DB 374 AGTTTGGAGACACATGCTCTCAAGAACATGATGGGACCCGAATTCGAACCTTAGTTT 433  
QY 477 CAAGAAATGGGACGATGAGTTTATTGAAGATGGGACCAAAATTCATGAAGGACCGCTTTT 536  
DB 434 CAGACAATTCGGAGCTGAGTATGTT---TTTATGCGCTGACTTTCATGAAGTTGGGTCTCT 490  
QY 537 ACCAAATTTGTCCCATAGAGATTACGAGCTGGCAAAATGTTGATAGGCAAGGTCAT 596

DB 491 ACCAGCTTCTCCAGTTGAGGATCTTGAATGGGATTAATTAATGAGGCCAGATCGT 550  
QY 597 TTTTTCACAGAGGATCTATCAAGAAAGAAAGTTTTCGCGAGGAAAGGATATGTTTCGGTGC 656  
DB 551 TATTTATTAACGATTTATCGAAGATGAAAACTTCTCGGATGAAGGATATGGTCTGTTTC 610  
QY 657 AACGAGTTTACGTAATGAGTAGTAGTGAAGACAAAGCATCCCTCGGATTTTCAATTCGTTGGA 716  
DB 611 CTCGAGTTTTCATAGTGTGTAAAGAGGACAAAGCAATTCAGAGAAAGCCGAGAGATGGA 670  
QY 717 TGATTGATAATTTCAACGCTCTCGAAAGTCTACGAGATCGATGCGCGAGATCACATGGTGA 776  
DB 671 TGATTGATAATTTTCCCGTGAATTTAGTGTGAGATGGAGAGACAGATCATATGCCAA 730  
QY 777 TGCTCTCCAAACCCCAAAACTCTTTGACTCTCTCTGCTTATTCGCCACCGATTATATGT 836  
DB 731 TGTCTGCAAGCCTCAGCAACTCAGTGATTAATCTTCTGAAATTCGCGACAAATTCGTTT 790  
QY 837 AA 838  
DB 791 AA 792  
RESULT 7  
AAC42344  
ID AAC42344 standard; DNA; 816 BP.  
XX AAC42344;  
XX AC AAC42344;  
XX 17-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 35196.  
XX KW Hybridisation assay; Genetic mapping; gene expression control;  
XX KW protein identification; signal transduction pathway; metabolic pathway;  
XX KW promoter; termination sequence; ss.  
XX OS Arabidopsis thaliana.  
XX PN EPI033405-A2.  
XX PD 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.



|                 |                 |                 |                |
|-----------------|-----------------|-----------------|----------------|
| PR 18-MAY-1999; | 99US-01334768P. | PR 27-JUL-1999; | 99US-0145919P. |
| PR 19-MAY-1999; | 99US-01334941P. | PR 28-JUL-1999; | 99US-0145951P. |
| PR 20-MAY-1999; | 99US-0135124P.  | PR 02-AUG-1999; | 99US-0146386P. |
| PR 21-MAY-1999; | 99US-0135353P.  | PR 02-AUG-1999; | 99US-0146388P. |
| PR 24-MAY-1999; | 99US-0135629P.  | PR 02-AUG-1999; | 99US-0146389P. |
| PR 25-MAY-1999; | 99US-0136021P.  | PR 03-AUG-1999; | 99US-0147038P. |
| PR 27-MAY-1999; | 99US-0136392P.  | PR 04-AUG-1999; | 99US-0147204P. |
| PR 28-MAY-1999; | 99US-0136782P.  | PR 04-AUG-1999; | 99US-0147302P. |
| PR 01-JUN-1999; | 99US-0137222P.  | PR 05-AUG-1999; | 99US-0147192P. |
| PR 03-JUN-1999; | 99US-0137528P.  | PR 05-AUG-1999; | 99US-0147260P. |
| PR 04-JUN-1999; | 99US-0137502P.  | PR 06-AUG-1999; | 99US-0147303P. |
| PR 07-JUN-1999; | 99US-0137724P.  | PR 06-AUG-1999; | 99US-0147416P. |
| PR 08-JUN-1999; | 99US-0138094P.  | PR 09-AUG-1999; | 99US-0147493P. |
| PR 10-JUN-1999; | 99US-0138540P.  | PR 09-AUG-1999; | 99US-0147935P. |
| PR 10-JUN-1999; | 99US-0138847P.  | PR 10-AUG-1999; | 99US-0148171P. |
| PR 14-JUN-1999; | 99US-0139119P.  | PR 11-AUG-1999; | 99US-0148319P. |
| PR 16-JUN-1999; | 99US-0139452P.  | PR 12-AUG-1999; | 99US-0148341P. |
| PR 16-JUN-1999; | 99US-0139453P.  | PR 13-AUG-1999; | 99US-0148565P. |
| PR 17-JUN-1999; | 99US-0139492P.  | PR 13-AUG-1999; | 99US-0148684P. |
| PR 18-JUN-1999; | 99US-0139454P.  | PR 16-AUG-1999; | 99US-0149368P. |
| PR 18-JUN-1999; | 99US-0139455P.  | PR 17-AUG-1999; | 99US-0149175P. |
| PR 18-JUN-1999; | 99US-0139456P.  | PR 18-AUG-1999; | 99US-0149426P. |
| PR 18-JUN-1999; | 99US-0139457P.  | PR 20-AUG-1999; | 99US-0149722P. |
| PR 18-JUN-1999; | 99US-0139458P.  | PR 20-AUG-1999; | 99US-0149723P. |
| PR 18-JUN-1999; | 99US-0139459P.  | PR 20-AUG-1999; | 99US-0149929P. |
| PR 18-JUN-1999; | 99US-0139460P.  | PR 23-AUG-1999; | 99US-0149902P. |
| PR 18-JUN-1999; | 99US-0139461P.  | PR 23-AUG-1999; | 99US-0149930P. |
| PR 18-JUN-1999; | 99US-0139462P.  | PR 25-AUG-1999; | 99US-0150566P. |
| PR 18-JUN-1999; | 99US-0139463P.  | PR 26-AUG-1999; | 99US-0150884P. |
| PR 18-JUN-1999; | 99US-0139750P.  | PR 27-AUG-1999; | 99US-0151065P. |
| PR 18-JUN-1999; | 99US-0139763P.  | PR 27-AUG-1999; | 99US-0151066P. |
| PR 21-JUN-1999; | 99US-0139917P.  | PR 27-AUG-1999; | 99US-0151080P. |
| PR 22-JUN-1999; | 99US-0139899P.  | PR 30-AUG-1999; | 99US-0151303P. |
| PR 23-JUN-1999; | 99US-0140353P.  | PR 31-AUG-1999; | 99US-0151438P. |
| PR 23-JUN-1999; | 99US-0140354P.  | PR 01-SEP-1999; | 99US-0151930P. |
| PR 24-JUN-1999; | 99US-0140695P.  | PR 07-SEP-1999; | 99US-0152363P. |
| PR 28-JUN-1999; | 99US-0140823P.  | PR 10-SEP-1999; | 99US-0153070P. |
| PR 29-JUN-1999; | 99US-0140991P.  | PR 13-SEP-1999; | 99US-0153758P. |
| PR 30-JUN-1999; | 99US-0141287P.  | PR 15-SEP-1999; | 99US-0154018P. |
| PR 01-JUL-1999; | 99US-0141842P.  | PR 16-SEP-1999; | 99US-0154039P. |
| PR 01-JUL-1999; | 99US-0142154P.  | PR 20-SEP-1999; | 99US-0154779P. |
| PR 02-JUL-1999; | 99US-0142055P.  | PR 22-SEP-1999; | 99US-0155139P. |
| PR 06-JUL-1999; | 99US-0142390P.  | PR 23-SEP-1999; | 99US-0155486P. |
| PR 08-JUL-1999; | 99US-0142803P.  | PR 24-SEP-1999; | 99US-0155659P. |
| PR 09-JUL-1999; | 99US-0142920P.  | PR 28-SEP-1999; | 99US-0156458P. |
| PR 12-JUL-1999; | 99US-0142977P.  | PR 29-SEP-1999; | 99US-0156596P. |
| PR 13-JUL-1999; | 99US-0143342P.  | PR 04-OCT-1999; | 99US-0157117P. |
| PR 14-JUL-1999; | 99US-0143624P.  | PR 05-OCT-1999; | 99US-0157753P. |
| PR 15-JUL-1999; | 99US-0144005P.  | PR 06-OCT-1999; | 99US-0157865P. |
| PR 16-JUL-1999; | 99US-0144085P.  | PR 07-OCT-1999; | 99US-0158029P. |
| PR 16-JUL-1999; | 99US-0144086P.  | PR 08-OCT-1999; | 99US-0158232P. |
| PR 19-JUL-1999; | 99US-0144325P.  | PR 12-OCT-1999; | 99US-0158369P. |
| PR 19-JUL-1999; | 99US-0144331P.  | PR 13-OCT-1999; | 99US-0159293P. |
| PR 19-JUL-1999; | 99US-0144332P.  | PR 13-OCT-1999; | 99US-0159294P. |
| PR 19-JUL-1999; | 99US-0144333P.  | PR 13-OCT-1999; | 99US-0159295P. |
| PR 19-JUL-1999; | 99US-0144334P.  | PR 14-OCT-1999; | 99US-0159329P. |
| PR 19-JUL-1999; | 99US-0144335P.  | PR 14-OCT-1999; | 99US-0159330P. |
| PR 20-JUL-1999; | 99US-0144352P.  | PR 14-OCT-1999; | 99US-0159331P. |
| PR 20-JUL-1999; | 99US-0144632P.  | PR 14-OCT-1999; | 99US-0159637P. |
| PR 20-JUL-1999; | 99US-0144684P.  | PR 14-OCT-1999; | 99US-0159638P. |
| PR 21-JUL-1999; | 99US-0144814P.  | PR 18-OCT-1999; | 99US-0159584P. |
| PR 21-JUL-1999; | 99US-0145086P.  | PR 21-OCT-1999; | 99US-0160741P. |
| PR 21-JUL-1999; | 99US-0145088P.  | PR 21-OCT-1999; | 99US-0160767P. |
| PR 22-JUL-1999; | 99US-0145087P.  | PR 21-OCT-1999; | 99US-0160768P. |
| PR 22-JUL-1999; | 99US-0145089P.  | PR 21-OCT-1999; | 99US-0160770P. |
| PR 22-JUL-1999; | 99US-0145192P.  | PR 21-OCT-1999; | 99US-0160814P. |
| PR 23-JUL-1999; | 99US-0145145P.  | PR 22-OCT-1999; | 99US-0160980P. |
| PR 23-JUL-1999; | 99US-0145218P.  | PR 22-OCT-1999; | 99US-0160981P. |
| PR 23-JUL-1999; | 99US-0145224P.  | PR 22-OCT-1999; | 99US-0160989P. |
| PR 26-JUL-1999; | 99US-0145276P.  | PR 25-OCT-1999; | 99US-0161404P. |
| PR 27-JUL-1999; | 99US-0145913P.  | PR 25-OCT-1999; | 99US-0161405P. |
| PR 27-JUL-1999; | 99US-0145918P.  | PR 25-OCT-1999; | 99US-0161406P. |

```
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      26.5%; Score 241.6; DB 3; Length 816;
Best Local Similarity 59.6%; Pred. No. 1.5e-67;
Matches 466; Conservative 0; Mismatches 304; Indels 12; Gaps 3;

QY 66 AGAGGAACATCAGTTCGGTTAGTTCACAAAGCTTATCATGGAGCCCTGGATCTGGTACA 125
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 38 AGAGGAACAACATTTGTTACTAGTACATGTTCTGTCGCATGGCGTGGTCTGGTACA 97

QY 126 AGCTCAAGCCCTCTCTTGAATCAGCGCGCACCGGTTACTCTCTCGAACTCGCGGCT 185
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 98 AGGTTAAGCGCTGTAGAGGCGGTGGGCCACCGGTAACTGCTGTGACTTAGCTGCT 157

QY 186 CCGGATCGAACCA---CGACCAATCCAGGCGGTTGAAACCGTTCGACGAATACTCCAAC 242
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 158 CCGGAATAGACACAACGAGGTGCTGACTGACATCCCAATGCCACATCGCAATACTCGGAGC 217

QY 243 CGTTGATCGAAACCTCAATCTCTTCAGAGAACGAAGAGTAATCTGTGTTGATTCAC 302
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 218 CATTGACGAAGTCTCTGACTCATTCGCCAATGATGAAGAAGTTGTCTGTTGTGTCACA 277

QY 303 GCTTCGGAGGCATCAACATCGCTCTCGCGCGACATATTTCCGGCGAAGATTAAAGTTTC 362
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 278 GCTTTGGTGGCTTGAACCTTAGCCATAGCCATGGAAGTTTCCCGAANAATCTGTGCG 337

QY 363 TTGTGTTCTCAACGCTTCTTTCGGCAGACCAACCGCTGCTCTCAAGTTCTTGAGACA 422
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 338 CTGTATCTTGACTGCTTTCATCGCGGACACCGCACTCAACATCTTCGTTCTTGACA 397

QY 423 AGTAT-----ATGGAGATGCTGGAGGTTTGGGAGATGTGAGTTTTCATCTCATGAA 476
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 398 AGTTTGGAAAGCAACATGCTCAAGAAGCATGTGATGGGCACCGAATTCGAACCTTATGTT 457

QY 477 CAAGAAATGGGACGATGAGTTTATTAAGATGGGACCAAAATTCATGAAGGCACGCTTTT 536
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 458 CAGACAAATCCGACTGAGTATGTT---TTTATGCTGACTTCATGAAGTTGGTCTCT 514

QY 537 ACCAAATTTGTCCTATAGAGATACAGCTGGCAAAAATGTTGCATAGGCAAGGTCAT 596
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 515 ACCAGCTTTCTCCAGTTGAGGATCTTGAATGGGATTACTTTTAATGAGGCCAGGATCGT 574

QY 597 TTTTTCACAGGATCTATCAAGAAGAAAGTTTATGCGAGGAGGATATGTTTCGGTGC 656
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 575 TATTTATTAACGATTTATCGAAGATGAANAATCTCTCGATGAAGATATGGGTCTGTTTC 634

QY 657 AACGAGTTTACGTAATGAGTAGTGAAGACAAAGCCATCCCTCGCGATTTCAATTCGTTGA 716
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 635 CTCGAGTTTTCATAGTGTGTAAAGGACAAAGCAATTCAGAGAAGCCAGAGATGGA 694

QY 717 TGAATGATAATTTCAACGCTCGAAGTCTACAGATCGATGGCGGAGATCATGTTGGA 776
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 695 TGATTTGATAATTTTCGGTGAATTTAGTGTAGATGGAGATGGAGGAGACAGATCATATGCCAA 754

QY 777 TGCTCTCCAAACCCCAAACTTTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 836
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 755 TGTCTGCAAGCCTCAGCAACTCAGTGATTTACTTCTGAAATTCGCGGCAAAATTCGTTT 814

QY 837 AA 838
DB |||
DB 815 AA 816
```

RESULT 8  
ADA26498  
ID ADA26498 standard; cDNA; 1079 BP.  
XX

```
AC ADA26498;
XX
DT 20-NOV-2003 (first entry)
XX
DE Tobacco salicylic acid binding protein 2 (SABP2) gene.
XX
KW salicylic acid-binding protein; SABP2; disease resistance; plant;
modulator; tobacco; ds; gene.
XX
OS Nicotiana tabacum.
XX
FH Key Location/Qualifiers
FT CDS 42..824
FT /*tag= a
FT /product= "salicylic acid binding protein SABP2"
XX
PN WO2003016551-A2.
XX
PD 27-FEB-2003.
XX
PF 16-AUG-2002; 2002WO-US026312.
XX
PR 16-AUG-2001; 2001US-0312863P.
XX
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX
PI Klessig DF, Kumar D;
XX
WPI; 2003-278578/27.
DR P-PSDB; ADA26499.
XX
PT New salicylic acid-binding protein 2 (SABP2) gene and protein, useful for
producing transgenic plants with increased resistance to disease, or for
screening SABP2 modulators that confer enhanced resistance of plants to
disease.
XX
PS Claim 1; Fig 4; 98pp; English.
XX
CC The invention relates to a novel isolated 1079 bp salicylic acid-binding
protein 2 (SABP2) nucleic acid molecule encoding a 261 amino acid
protein, sequences that specifically hybridize with the nucleic acid, the
complement of the nucleic acid or a natural allelic variant of the
sequence. The nucleic acid, SABP2 or a SABP2 homologue is useful for
conferring increased resistance to disease in plants, or for screening
modulators of SABP2, which confer increased or enhanced resistance of
plants to disease. The SABP2 or SABP2 homologue nucleic acid is also
useful for producing transgenic plants with increased resistance to
disease. The SABP2 nucleic acid, SABP2, or its homologue, is also useful
for studying the molecular mechanisms responsible for the modulation of
salicylic acid-mediated disease resistance in plants. This sequence
represents the gene encoding the SABP2 protein from tobacco plants.
XX
SQ Sequence 1079 BP; 307 A; 204 C; 234 G; 334 T; 0 U; 0 Other;
Query Match 25.5%; Score 232.6; DB 8; Length 1079;
Best Local Similarity 56.3%; Pred. No. 1.5e-64;
Matches 458; Conservative 0; Mismatches 349; Indels 6; Gaps 1;
QY 33 AACCAAAAAGAAAACGCTCTACGGAGACATATGGAGAGGAAACATCATCTTCGTGTAGTTC 92
DB 13 AAAAGAACTTAAACAGGCATAAAATTCAAATGAAGGAAGAAACACATTTGTTTTAGTAC 72
QY 93 ACAACGCTTATCATGGAGCCTGGATCTGGTACAGCTCAAGCCCTCTCTGAATCAGCGG 152
DB 73 ATGGTGCATGCCATGGAGGTTGGAGTTGGTCAAGCTAAAGCCACTGCTAGAGCTGCAG 132
QY 153 GCCACCGCTTACTGTCTCGAACTCGCGGCTCGGGGATCGACCCAGCAACCAATCCAGG 212
DB 133 GCCATAGGTTACAGCCCTTGATTTAGCAGCTTCTGGCAGTATTTGAGAAAATAGAGG 192
QY 213 CCGTTGAAACCGTTCGACGAATACTCCAAACCGTTGATCGAAACCCCTCAAAATCTCTTCAG 272
DB 193 AGCTTCGCACACATTTATGATTATATCTTTGCCATTTGATGAGTTGATGGAATCTCTTTTCAG 252
```



Db 753 TGTATGCGGAAGGATCTTGTGTACCCGGAAGATTACAGCGATCGATGATCAGCACTT 812  
Qy 730 CAACGTCTCGAAAGTCTACGAGATCGATGGCGGAGATCACATGGTGATGCTCTCCAAACC 789  
Db 813 TCCCCCAAAAGAAGTAATGGAGATCAAGAGCGCAGATCATATGCCAATGTTCTCCAAGCC 872  
Qy 790 CCAAAAACCTTTTGACTCTCTCTCTGCTATTCGCCACCGATATATATGTA 838  
Db 873 TCAACAACATATGTCTCTCTCTGAGATTGCAATAAATATGCTAA 921

RESULT 10  
AAT36351  
ID AAT36351 standard; cDNA; 1091 BP.  
XX  
AC AAT36351;  
XX 29-OCT-1997 (first entry)  
XX Hevea brasiliensis S-hydroxynitrilase cDNA.  
DE S-hydroxynitrilase; chiral synthesis; aliphatic; aromatic; S-cyanohydrin;  
KW recombinaant; ss.  
KW  
XX Hevea brasiliensis.  
XX  
FH Key Location/Qualifiers  
FT CDS 57..830  
FT FT /\*tag= a  
FT FT /product= "S-hydroxynitrilase"  
XX  
PN W09703204-A2.  
XX  
XX 30-JAN-1997.  
XX  
XX 10-JUL-1996; 96WO-EP003010.  
XX  
XX 12-JUL-1995; 95AT-00001182.  
XX  
XX (STAM ) DSM CHEMIE LINZ GMBH.  
XX  
XX Hasslacher M, Schall M, Schwab H, Hayn EM, Kohlwein S, Griengl H;  
XX  
XX WPI; 1997-119058/11.  
DR P-PSDB; AAW10593.  
XX  
XX Hevea brasiliensis S-hydroxynitrilase - and recombinant equivalent of  
PT high specific activity, for chiral synthesis of cyano:hydrin(s).  
XX  
XX Claim 2; Page 2-3; 14pp; German.

XX The present sequence encodes the Hevea brasiliensis S-hydroxynitrilase (S  
CC -HN), which can be used for the chiral synthesis of aliphatic or aromatic  
CC S-cyanohydrins from HCN, or its donor, and ketone. Recombinant S-HN has  
CC higher specific activity than the native enzyme. H. brasiliensis leaves  
CC were homogenised in cold pH 6.5 K phosphate buffer, and the extract  
CC subjected to sequential chromatography on QAE-Sephacrose, Phenyl-Sepharose  
CC and BioGel 150 to recover a protein with a specific activity of 19 IU/mg.  
CC A cDNA library was prepared from young H. brasiliensis leaves in Zap  
CC phage, and screened with polyclonal rabbit antiserum raised against the  
CC purified enzyme. The insert in one positive clone was isolated, and  
CC ligated into pBluescript to form pHNL-100. This plasmid, in E. coli SOUR,  
CC expressed a fusion protein with lacZ that was immunoreactive. Sequencing  
CC showed that the insert in this plasmid was not complete, and the missing  
CC 5'-end generated by 2 stage PCR, resulting in plasmid pHNL-101,  
CC containing the entire coding sequence

XX Sequence 1091 BP; 326 A; 196 C; 239 G; 330 T; 0 U; 0 Other;  
SQ  
Query Match 24.3%; Score 221.6; DB 2; Length 1091;  
Best Local Similarity 55.3%; Pred. No. 6.1e-61;  
Matches 431; Conservative 0; Mismatches 349; Indels 0; Gaps 0;

Qy 76 TCATCTTGTTAGTTTACAAAGCTTATCATGGAGCCTGGATCTGGTACAGCTCAAGCC 135  
Db 68 TCATTTTGTCTTATTCATACCATATGCCGGTGCATGGATTTGGCAAGCTCAAAACC 127  
Qy 136 CCTCTTGAATCAGCCGCCACCGCGTTTACTGTGTGAACTCGCCGCTCCGGGATCGA 195  
Db 128 CCTCTTGAGGCACTTGGCCACAAGGTTACTGCATGGACCTTGACGCAAGCGGGTTGA 187  
Qy 196 CCCAGCAACATCCAGCCGTTGAAACCGTGCAGCAATATCCAAACCGTTGTATCGAAAC 255  
Db 188 CCCAAGGCANAATTGAGGAGATTGGCTCATTTGATGAGTATTTCTGAACCTTGTGTGACGTT 247  
Qy 256 CCTCAATCTCTCCAGAGAACGAGAGTAATTTCTGTTGCGATTTCAGCTTCGAGGCGAT 315  
Db 248 CTTGAGGCACTCCCTCCAGGGGAAAGGTAATTTCTGTTGGCGAGAGCTGTGAGGACT 307  
Qy 316 CAACATCGCTCTCGCCGCCGACATATTTCCCGCGCAAGATTAAAGTTCTTGTGTCTCTCAA 375  
Db 308 CAATATAGCAATTGCTGCTGATTAATACTGTGAAAGATTTCAGCTGCTGTTTCCACAA 367  
Qy 376 CGCTCTTCTGCCCAGACAAACCGTGCCTTCTCAGCTTCTGGACAAGTATATGGAGAT 435  
Db 368 TTCAGTATTGCCAGACACCGGAGCACTGCCATCTTACGTCGTGGATAAGCTCATGGAGGT 427  
Qy 436 GCCTGGAGGTTTGGGAGATTGTGAGTTTTCATCTCATGAAACAAGAAATGGAGCATGAG 495  
Db 428 GTTTCGCCGACTGGAAAGACACACGTAATTTTACGTACACTAAAGATGCCAAGGAGATAAC 487  
Qy 496 TTATTTGAAGATGGGACCAAAATTCATGAAGGCGACGCTCTTTTACCAAAATTTGTCCCATAGA 555  
Db 488 TGGATTGAACTGGGCTTCAGCTTCTGAGGGAATAATTTATATACCTTTGCGGTCTCTGA 547  
Qy 556 GGATTACGAGCTGGCAAAATTTGTCATAGCAAGGAGGTCATTTTTCACAGAGGATCTATC 615  
Db 548 GGAATATGAACCTGGCGAAGATGTTTCAAGGAAGGAGTCATTTTCAAAAATATTTTAGC 607  
Qy 616 AAGAAAGAAAGTTTACGAGGAAGGATATGTTTCGTTGCGCAACAGTTTACGTAATGAG 675  
Db 608 TAAGCGACCATTTCTTCACTAAGGAAGGTTACGGATCGAATTAAGAAAATTTATGTGTGAC 667  
Qy 676 TAGTGAACACAAAGCCATCCCTCGGATTTTCATTTCTGTTGGATGATGATAATTTCAACGT 735  
Db 668 CGACCAAGCAATATTTTACCTGAATTTTCAACTCTGGCAAAATAGAAAATCTATAAAC 727  
Qy 736 CTCGAAAGCTTACGAGATCGATGGCGGAGATCATGATGATGCTCTCCAAACCCCAAAA 795  
Db 728 AGACAAGGTTTATAAGGTCGAAGGTGGAGATCATANAATTGCGCTTACAAAGACTAAGGA 787  
Qy 796 ACTCTTTGACTCTCTCTGCTATTGCCACCGATTATATGTAATAATCTTAAAGTCGGTTT 855  
Db 788 GATCGCTGAAATTTCTCAAGAGGTGGCTGATACCTATTAATTGACTTCTTTGAGGGCTTTT 847

RESULT 11  
AAT86824  
ID AAT86824 standard; cDNA; 817 BP.  
XX  
AC AAT86824;  
XX 09-JAN-1998 (first entry)  
XX Hevea brasiliensis (S)-hydroxynitrilase cDNA.  
DE (S)-hydroxynitrilase; production; (S)-cyanohydrin; aldehyde;  
KW hydrogen cyanide; recombinaant; ss.  
XX  
XX Hevea brasiliensis.  
XX  
FH Key Location/Qualifiers  
FT CDS 44..817  
FT FT /\*tag= a  
FT FT /product= "(S)-hydroxynitrilase"